

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 10:21:53 ; Search time 1822.67 Seconds
(without alignments)
3926.589 Million cell updates/sec

Title: US-09-829-124-1
Perfect score: 342
Sequence: 1 atggactctatcggaacaa.....gcagctggcggaacgcc 342

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank.*
1: gb_ba.*
2: gb_hgt.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_hum.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htgo_inv.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES			
Result No.	Score	Query Match Length	Description

1	342	100.0	342	6	AX256351
2	49.2	14.4	10096	1	AF026197
3	48.6	14.2	767	3	EIMMAX
4	48	14.0	19304	1	AB045311
5	45	13.2	162869	9	AC090645
6	43.8	12.8	207050	1	AL646063
7	43.2	12.6	190770	2	AC055876
8	43	12.6	178984	2	AC084023
9	42.4	12.4	159971	2	AC091951
10	42.2	12.3	13409	1	AE004466
11	42	12.3	9921	3	AF325500
12	42	12.3	142277	2	AC016204
13	41.4	12.1	175296	2	AC106674
14	41.2	12.0	204152	2	AC092203
15	40.8	11.9	272545	2	AC090533
16	40.6	11.9	5719	10	MMCA7S1
17	40.6	11.9	197265	10	AL603702
18	40.6	11.9	320096	2	AL603889
19	40.2	11.8	14034	1	AE005694
20	40.2	11.8	215046	2	AC011767
21	40.2	11.8	277577	2	AC095395
22	40	11.7	72187	2	AC013433
23	40	11.7	199508	2	AC098555
24	39.8	11.6	10565	1	AE004576
25	39.8	11.6	10729	1	AB012767
26	39.8	11.6	11482	1	AE001876
27	39.8	11.6	40510	3	AF067607
28	39.8	11.6	160309	2	AC099443
29	39.8	11.6	236962	2	AL606742
30	39.4	11.5	209128	5	AB073377
31	39	11.4	1212	3	PFACSL
32	39	11.4	110000	2	LMFLCHR34_01
33	39	11.4	190803	9	AL355490
34	39	11.4	204050	1	AL646073
35	38.6	11.3	5174	9	HSCOLA43
36	38.6	11.3	110000	2	LMFLCHR34_11
37	38.4	11.2	6814	5	AF297658
38	38.4	11.2	80009	3	AF226688
39	38.4	11.2	81053	3	AC005762
40	38.4	11.2	89428	2	AC017242
41	38.4	11.2	208907	2	AC106604
42	38.4	11.2	215050	1	AL646057
43	38.4	11.2	240021	2	AC099436
44	38.4	11.2	283698	2	AC099294
45	38.4	11.2	302473	3	AE003589

ALIGNMENTS

RESULT 1	AX256351	342 bp	DNA	linear	PAT 10-OCT-2001
LOCUS	Sequence 8 from Patent WO0170988.				
DEFINITION	AX256351				
ACCESSION	AX256351.1	GI:16075202			
VERSION					
KEYWORDS	Xanthomonas campestris pv. pelargonii.				
SOURCE	Xanthomonas campestris pv. pelargonii.				
ORGANISM	Xanthomonas campestris pv. pelargonii.				
REFERENCE	1 (bases 1 to 342)				
AUTHORS	Song,X., Fan,H. and Wei,Z.M.				
TITLE	Receptors for hypersensitive response elicitors and uses thereof				
JOURNAL	Patent: WO 0170988-A 8 27-SEP-2001;				
EDEN	Eden Bioscience Corporation (US)				
FEATURES	Location/Qualifiers				
source	1. 342				
BASE COUNT	83 a 94 c 109 g 56 t				
ORIGIN	/organism="Xanthomonas campestris pv. pelargonii"				
	/db_xref="taxon:91612"				

Query Match 100.0%; Score 342; DB 6; Length 342;
Best Local Similarity 100.0%; Pred. No. 1.5e-62;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 atgactctatcgaaacatttccgaataatcgcaacctgcagacgagtcggtcggc 60
Db 1 ATGACTCTATCGGAAACAATTTCGAATATCGCAACCTGCAGACGATCGGATCGG 60
Qy 61 cctcagcaacacagagactcagcagcagtcgcttcgctcgcgctcgcgagcagctg 120
Db 61 CCTCAGCAACACAGAGACTCAGCAGCAGTCGCTTCGGCTCGGCTCGGAGCAGCTG 120
Qy 121 gatcagttgctcgccatgttcacatgatgatctgcaacagacccagcagcagatga 180
Db 121 GATCAGTTGCTCGCATGTTTCATCATGATGATGCTGCAACAGACCCAGGCGGATGCA 180
Qy 181 aatcaggagtggtgcaacacacagcagcagtggtcaacaggagcctgagtcgcttg 240
Db 181 AATCAGGAGTGTGGCAACGAACAACCGCAGACGCTCAACAGGAAGCCTGAGTCGCTG 240
Qy 241 accagatgctgatcagatcgatcgatcgatcgatcgatcgatcgatcgatcgatcg 300
Db 241 ACCAGATGCTGATCGAGATCGTATGATGATGATGATGATGATGATGATGATGATG 300
Qy 301 ggcgggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggtggt 342
Db 301 GCGGGTGGCGGTGCGGTGCAACAGCAGCCTGGCGGCAACGCC 342

RESULT 2

AF026197 10096 bp DNA linear BCT 30-MAR-2000
LOCUS Xanthomonas oryzae pv. oryzae hrp gene cluster, partial sequence.
DEFINITION AF026197 AF232057
ACCESSION AF026197.2 GI:7350908
VERSION
KEYWORDS
SOURCE Xanthomonas oryzae pv. oryzae.
ORGANISM Xanthomonas oryzae pv. oryzae
Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
Xanthomonas.
REFERENCE 1 (bases 1 to 10096)
AUTHORS Zhu, W., Magbanua, M.M. and White, F.F.
TITLE Identification of two novel hrp-associated genes in the hrp gene cluster of xanthomonas oryzae pv. oryzae
J. Bacteriol. 182 (7), 1844-1853 (2000)
MEDLINE 20179797
PUBMED 10714988
REFERENCE 2 (bases 891 to 2420)
AUTHORS Zhu, W. and White, F.F.
TITLE HrpXo is a regulator of hrp gene expression in Xanthomonas oryzae pv. oryzae
Unpublished
JOURNAL
REFERENCE 3 (bases 891 to 2420)
AUTHORS Zhu, W. and White, F.F.
TITLE Direct Submission
Submitted (22-SEP-1997) Plant Pathology, Kansas State University, Throckmorton Hall, Manhattan, KS 66506, USA
JOURNAL
REFERENCE 4 (bases 1 to 10096)
AUTHORS White, F.F., Zhu, W. and Magbanua, M.M.
TITLE Direct Submission
Submitted (08-FEB-2000) Plant Pathology, Kansas State University, 4024 Throckmorton Hall, Manhattan, KS 66506, USA
JOURNAL
REMARK Sequence update by submitter
COMMENT On Mar 30, 2000 this sequence version replaced gi:4003499.
FEATURES
1..10096
source
/organism="Xanthomonas oryzae pv. oryzae"
/strain="PX086"
/db_xref="taxon:64187"
/note="pathovar: oryzae"
misc_feature
1..6
/note="adjoins with sequence presented in GenBank

Accession Number AF232058"
complement(129..908)
/gene="hpa2"
complement(129..617)
/gene="hpa2"
/note="lysozyme related protein; hrp-associated protein"
/codon_start=1
/transl_table=11
/evidence="not experimental"
/product="Hpa2 precursor"
/protein_id="AAF61278.1"
/db_xref="GI:7350910"
/translation="MLFAALACAAAFARACDFEAAAGYOHVPWVLRATAWDSRGR
ADAIHRNNGTVDYGRMQINSIHLRLFGYISKEALMQPCSVYVAARLREMTNKY
GNTAAVAGVYHSETPGERDKYAHAIHSILRLGLVYVNRASDPAGMRGAAMRHLHQ
VA"
complement(570..617)
/gene="hpa2"
complement(884..908)
/gene="hpa2"
/note="variant plant-inducible promoter, Pip-1"
976..1000
/gene="hpa1"
/note="variant plant-inducible promoter, Pip-2"
976..1567
/gene="hpa1"
1136..1567
/gene="hpa1"
/note="Orf432; HrpXo; similar to Xanthomonas oryzae pv.
oryzae PopA; hrp associated protein; glycine-rich protein
encoded by hrp region"
/codon_start=1
/transl_table=11
/product="Hpa1"
/protein_id="AAC95121.1"
/db_xref="GI:4003500"
/translation="MNSLNTQFGSTNLQVGSQDTTFGSNGGNOGISEKOLDQL
COLISALQSSNAEKGKGGGNGGQSQGQNGPSPFTOMLHIVGEILQA
QNGGAGGGGFGGFGGFGGFGGDLGLTNLSSGSASMQ"
complement(2306..4123)
/gene="hrpA"
complement(2306..4123)
/gene="hrpA"
/note="Hrcc related"
/codon_start=1
/transl_table=11
/product="Hrpa"
/protein_id="AAC95122.2"
/db_xref="GI:7350909"
/translation="MAPACTTTHRQAPLFAVLLSLFLSPQADAAQVPHSRTFK
YVADNKLKEVLRLDSASQSIATWISPEVTGLSGKFETSPQKFLDLDLAATYGFVY
DGAMLRINGANESKATLSLGTASTKSLRDLARLDDSPRYDEAAHVAVVSGP
PGYVDTVSAIAKQVEGARORDATEVQVQALHYAQRADHTTRIGQDVQIPGASLLR
SIYGARGSAVPIAAGFANFGRVQPIGGSSNTFGNAOCGGGASGTGLGSPANFGG
ASPDPMVPSPLPGSAAAGSPASVMPSELKSRDESNPIDAGGAELASDAPVIEA
DPTNAILIRDRPERMQSYGLTIQDLNRPKLLQDITAEIRDMAGMDLGVDRFHS
NKAKTVSSPOVATLDNVEAMDKQAFVRSVGSYASADYLNLSAGVSLRVLPSVPGS
PNGOMLRVRIEDGOLGNTVDGIPVITSSEIKTOAFVNEGQSLIAGYAYDADETDL
NAVGLSKIPLLGNLFKRRKSGSRMQRLLFTLTPHVSF"
complement(4206..5036)
/gene="hrpB"
complement(4206..9926)
/gene="hrpB operon"
complement(4206..5036)
/gene="hrpB"
/codon_start=1
/transl_table=11
/product="HrpB"
/protein_id="AAF61279.1"
/db_xref="GI:7350911"
/translation="MNDVTDALLSSQGVSLTLLALCGVRVFWFVFLPATQDSL
PGIARNGVIYVLSFATYAGQPADALAKTQTVGLGVGWFKEAFIGLLIGFAAFWIA


```

/protein_id="AAG03667.1"
/db_xref="GI:9946121"
/translation="MSVDLHALPLIALTELLAGFVKGVVGIPLPTVSVGLGLAMP
MQAAALLVPSWYNLWQLACGRFGLMKRLAGLLGVVGVVGLLGLGADPQQA
TGVLGLAYALGLGLAAAPLPHPARHEPWAGPLVGLTALGTAVTGCVFVPAVPL
ALGQDELVOALGLSASTLALAVTLGVHGDLEPQMLGASLLTLVPAALGMLLQ
WLRORISATLFRRCFFVGLLLGLADLAWRGFH"
gene
CDS
  /gene="PA0279"
  /note="1661"
  /gene="PA0279"
  /codon_start=1
  /transl_table=11
  /product="probable transcriptional regulator"
  /protein_id="AAG03668.1"
  /db_xref="GI:9946122"
  /translation="MQFNPGFSSLAALLADPPGQAQMIWALMDGSARPASELALLAGVS
  PSSAGHLRGVLEAGVGRNFRFYLAGVEQVGAELASALQPPQRSLPP
  SHGTPSALREARTCYDHLGELAGVGFARMLDAGNTEQEGRTLRLGATGEAGLAGIGI
  DLAEVRRRRQFACACPDNSRKPRLGGALGAALLEACLRQGLRPQDGSRALQVSSK
  GRAGLRGAERTAG"
  complement(1674..2663)
  /gene="cysA"
  /note="PA0280"
  complement(1674..2663)
  /gene="cysA"
  /codon_start=1
  /transl_table=11
  /product="sulfate transport protein CysA"
  /protein_id="AAG03669.1"
  /db_xref="GI:9946123"
  /translation="MSIERNVSKNFNAFKALDDINLIQSGLVALLGPSGGKTKL
  LRLIAGLEPPDAGNVFHGEDVSQHDVRNNGVFVQHFALFRHMTVFONVAFGLRMK
  PAGERGESAIKAKVHELLNMVLQDLADRYPEQLSGQQRQIALARALAVPEPKILL
  DEFFGALDAKVELRWLRARHEEINTSVFTHDQEEAMEVADRIVVMNGVIEQI
  GSPGVEYSPASDFVYHFLGNDNLQNDHLLFRPEHVSLSRSASAVHRAEVRDI
  RPLGATRVTLAVDQODELIEAEVVKDHDLSLAGLARGETLYFKPAQFPAVL"
  complement(2667..3536)
  /gene="cysW"
  /note="PA0281"
  complement(2667..3536)
  /gene="cysW"
  /codon_start=1
  /transl_table=11
  /product="sulfate transport protein CysW"
  /protein_id="AAG03670.1"
  /db_xref="GI:9946124"
  /translation="MSSASISAAATASASRRGNAGGRALALIVLAWLVEALFLLLPLVY
  LSEALKQGGTFEFAILEPDALAALKLTLIAVASVPLNLLEGVAANACVSKFERGK
  SLTILDLIPSVSVIAGLIYVLLFGAGYFGEWLSHDHDIQIVFAVPGIVLITLFTV
  VFFVARELIPMQEGTQEEEAARLLGANGWQMFHWTLPLNIKWGLIYGVILCTARAM
  GEFGAVSVVSGHIGVNTLPLHVEILYNEYNHVAASVAVSALLLMLVILLKQWSE
  SRMSRLKSNADDE"
  complement(3547..4365)
  /gene="cysT"
  /note="PA0282"
  complement(3547..4365)
  /gene="cysT"
  /codon_start=1
  /transl_table=11
  /product="sulfate transport protein CysT"
  /protein_id="AAG03671.1"
  /db_xref="GI:9946125"
  /translation="MSRRISVIPFGTLTGTYLIVYLSLLVLIPLIGAMFLKTTQLSWE
  QFWAISAPRYLAALKSFGTALAARAVNLGTLIAWLVRYEFPGRKIIDAMIDL
  FALPTAVAGIATLALYAPAGLVQGFASDLGFKIATYPTGLITLALFTVLPFVTVQ
  VLADIPKEVEEAACIARPLQVFRHILVALLPAWLGFEALAFARGVGEYSVIFA
  GNMPMKTEILPLILMWKLDQDYTGATAIGVIMLVVSWFILLILLINLQRIETP"
  complement(4527..5525)
  /gene="sbp"
  /note="PA0283"
  complement(4527..5525)
  /gene="sbp"

```

```

/codon_start=1
/transl_table=11
/product="sulfate-binding protein precursor"
/protein_id="AAG03672.1"
/db_xref="GI:9946126"
/translation="MSLRREALAALASALYSGPVAAATOLLNVSYDPTRELYQAYNNA
FIKHKAQGGEDLTVQSHGSGKQARAVIDGLKADVYTLALAGDIDELHKLGLLPA
DWOARLPDNPSTPTITVFLVRKGNPKGIDKMDGLTKEGVEVITPNPKTSGGARWNL
LAAMAKQYGSDEKADYQALYKHVPDLDTGARGSTITFVNNQIGDVLAWENAF
LAKKEQGENFEIVPSISILAEPPVAVDVVKVEKGTGRVAAEAYLQYLYSEGORIA
AQNFYRPNOKVAAEFATQPKNLNLVTVDSDFCGWKTAQPKFFENDGGIFDQIYQAO"
complement(5702..5884)
/gene="PA0284"
complement(5702..5884)
/gene="PA0284"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="AAG03673.1"
/db_xref="GI:9946127"
/translation="MSATLRSLDQDEATILREIOSALRGLRFGAVEITVHNGQVVQI
ERKEKFLQOPAVKQA"
complement(6048..8330)
/gene="PA0285"
complement(6048..8330)
/gene="PA0285"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAG03674.1"
/db_xref="GI:9946128"
/translation="MSPRLSARLSLAYLAIAIAMAFLSDHLLQNVQDQDLVFWMT
AKRFFPAATGLLLYLVQFRKKAQYDELHGSQRNRALAEAVRDGLWDWDLVTD
RMFVSPGYAALIGLAEELGDPETWKRLHPEEYATVLEAHRNHLQGLTNDLHLYR
LRHKDQYWIHSRGRVLRDALGKPLHYTGVAEDITLRLKEDHLQQAAYEDSTREG
VLYTDAQAVIVHNPSPERITGTRESVDLGTIPAILRSRQDQAFQRLWLAUREODV
WSEGINRRKSGEITYPWLHIRAVNDQGLTHYGVFSDLSIKRSENELEDFLAHD
SUTGLPNRVLRLERIBQALENGKDRTVAGALLIDLDFKHINDSLGHTTGMLLKEV
SKRLQHLDERCLLSRLGDEFAILVENDDEPAVARLSQRLDGNAPFDIHCOPIYI
SASLGVSLYPEDASVDVHLMADALFOAKDSGRNAYAFYTRVLTARARAVOVESA
LRHALEHDELRYHOPVHDLASGRIVGESLYRQHPERGLYPPGFEVPAECCGLIA
ALDNVYLKACRQMRQQRQVELEFVAVNSRFLRGGLEERIANALESCLEPRY
LELEVTESAVMEDFEQSLNLLRLILGVNLAIIDDFGTQSYSLMRKLRLPVLHKLIDQ
GFVAGLPGAVDDAAATARAIVALAQAQSLRVVRAEGIEHQDQALFLREHGDGFGQYWG
RPQAEALRFDLPVVALPED"
8502..9680
/gene="PA0286"
8502..9680
/gene="PA0286"
/codon_start=1
/transl_table=11
/product="probable fatty acid desaturase"
/protein_id="AAG03675.1"
/db_xref="GI:9946129"
/translation="MWFNGFDLSPWQLVAVTLVLTHTITVSVTVYLHRYSAHRALDL
HPALQHFREFWLMTTGMREWTATHRKHAKCETDEPSPHKGFLTVLRAGAEL
YKAEARNQDITLRIYKNGCPDDWIERNLSRPIGGVILMALIDLALFGALGTVWAVO
MVIIPFAAGVYNGIYGHAGVYRNRDAATNLVPMGLITGGEELNHHHTYPSNAKL
QYKWEFDLGNWIKLFSFLRLARVAPVIAHRYEGKHSMDMDAMAILNNRFOIMA
SVYKLVIAPLVQEVAKADESVRLFRRAKRLSRETSLLQDRHHVHDSMLHSAOL
KVYIEKRLAQOIWAKTSANGHDMIAAIKDWVHEAASGIOSLSKEFAAQLKTYSLRPA"
9911..11296
/gene="PA0287"
9911..11296
/gene="PA0287"
/codon_start=1
/transl_table=11
/product="probable sodium:solute symporter"
/protein_id="AAG03676.1"
/db_xref="GI:9946130"
/translation="MLLDYLLIMLYALAMLGIVGWGRKAKSQSDFLVAGRRLGPLY
LGTMAAVVILGASTIGTVKLYGQFGLSGLVFLMGLGIIIVLSLVFSRQIARLVFTY

```


Quiles, M., Ren, Y., Rives, M., Rojas, A., Rojibokan, I., Rolfe, M.,
 Ruiz, S., Savary, G., Scherer, S., Scott, G., Shen, H., Shoohtari, N.,
 Sisson, I., Sodergren, E., Sonalike, T., Sparks, A., Stanley, H.,
 Stone, H., Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K.,
 Tang, H., Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N.,
 Thomas, S., Usmani, K., Vasquez, L., Vera, V., Villalon, D., Vinson, R.,
 Wall, R., Wang, S., Ward-Moore, S., Warren, R., Washington, C.,
 Watlington, S., Williams, G., Williamson, A., Wleczyk, R., Wooden, S.,
 Worley, K., Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
 Weinstein, G., and Gibbs, R.

Direct Submission

Unpublished
 2 (bases 1 to 175296)

Worley, K.C.

Direct Submission

Submitted (12-JAN-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA

COMMENT

----- Genome Center

Center: Baylor College of Medicine

Center code: BCM

Web site: <http://www.hgsc.bcm.tmc.edu/>

Contact: hgsc-help@bcm.tmc.edu

----- Project Information

Center project name: G10J

Center clone name: CH230-108D18

----- Summary Statistics

Assembly program: Phrap; version 0.990329first call to

findPhrapList

Consensus quality: 148335 bases at least Q40
 Consensus quality: 155812 bases at least Q30
 Consensus quality: 162105 bases at least Q20
 Estimated insert size: 152728; sum-of-contigs estimation
 Quality coverage: 0x in Q20 bases; agarose-fp estimation
 Quality coverage: 2.4x in Q20 bases; sum-of-contigs estimation

 * NOTE: Estimated insert size may differ from sequence length
 * (see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html).
 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 56 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 14942: contig of 14942 bp in length
 * 14943 15042: gap of unknown length
 * 15043 26820: contig of 11778 bp in length
 * 26821 26920: gap of unknown length
 * 26921 31321: contig of 4401 bp in length
 * 31322 31421: gap of unknown length
 * 31422 37389: contig of 5968 bp in length
 * 37390 37489: gap of unknown length
 * 37490 42498: contig of 5009 bp in length
 * 42499 42598: gap of unknown length
 * 42599 46308: contig of 3710 bp in length
 * 46309 46408: gap of unknown length
 * 46409 52395: contig of 5987 bp in length
 * 52396 52495: gap of unknown length
 * 52496 59139: contig of 6644 bp in length
 * 59140 59239: gap of unknown length
 * 59240 64327: contig of 5088 bp in length
 * 64328 64427: gap of unknown length
 * 64428 68633: contig of 4206 bp in length
 * 68634 68733: gap of unknown length
 * 68734 73329: contig of 4595 bp in length
 * 73329 73428: gap of unknown length
 * 73429 77037: contig of 3609 bp in length
 * 77038 77137: gap of unknown length
 * 77138 81483: contig of 4346 bp in length
 * 81484 81584: gap of unknown length
 * 81584 85544: contig of 3961 bp in length

85545 85644: gap of unknown length
 * 85645 89737: contig of 4093 bp in length
 * 89738 89837: gap of unknown length
 * 89838 94559: contig of 4722 bp in length
 * 94560 94659: gap of unknown length
 * 94660 96754: contig of 2095 bp in length
 * 96755 96854: gap of unknown length
 * 96855 98841: contig of 2987 bp in length
 * 98842 99941: gap of unknown length
 * 99942 103396: contig of 3455 bp in length
 * 103397 103496: gap of unknown length
 * 103497 106178: contig of 2682 bp in length
 * 106179 106278: gap of unknown length
 * 106279 109015: contig of 2737 bp in length
 * 109016 109115: gap of unknown length
 * 109116 111536: contig of 2421 bp in length
 * 111537 111636: gap of unknown length
 * 111637 114747: contig of 3111 bp in length
 * 114748 114847: gap of unknown length
 * 114848 117377: contig of 2530 bp in length
 * 117378 117477: gap of unknown length
 * 117478 120113: contig of 2636 bp in length
 * 120114 120213: gap of unknown length
 * 120214 122860: contig of 2647 bp in length
 * 122861 122960: gap of unknown length
 * 122961 125473: contig of 2513 bp in length
 * 125474 125573: gap of unknown length
 * 125574 128271: contig of 2898 bp in length
 * 128272 128371: gap of unknown length
 * 128372 131331: contig of 2960 bp in length
 * 131332 131431: gap of unknown length
 * 131432 133677: contig of 2246 bp in length
 * 133678 133777: gap of unknown length
 * 133778 135703: contig of 1926 bp in length
 * 135704 135803: gap of unknown length
 * 135804 137886: contig of 1983 bp in length
 * 137887 137886: gap of unknown length
 * 137887 140617: contig of 2731 bp in length
 * 140618 140717: gap of unknown length
 * 140718 141934: contig of 1217 bp in length
 * 141935 142034: gap of unknown length
 * 142035 144079: contig of 2045 bp in length
 * 144080 144179: gap of unknown length
 * 144180 146430: contig of 2251 bp in length
 * 146431 146530: gap of unknown length
 * 146531 148261: contig of 1731 bp in length
 * 148262 148361: gap of unknown length
 * 148362 150017: contig of 1656 bp in length
 * 150018 150117: gap of unknown length
 * 150118 151592: contig of 1475 bp in length
 * 151593 151692: gap of unknown length
 * 151693 153207: contig of 1515 bp in length
 * 153208 153307: gap of unknown length
 * 153308 154357: contig of 1050 bp in length
 * 154358 154457: gap of unknown length
 * 154458 156066: contig of 1609 bp in length
 * 156067 156166: gap of unknown length
 * 156167 157545: contig of 1379 bp in length
 * 157546 157645: gap of unknown length
 * 157646 159211: contig of 1566 bp in length
 * 159212 159311: gap of unknown length
 * 159312 160381: contig of 1070 bp in length
 * 160382 160481: gap of unknown length
 * 160482 161912: contig of 1431 bp in length
 * 161913 162012: gap of unknown length
 * 162013 163616: contig of 1604 bp in length
 * 163617 163716: gap of unknown length
 * 163717 164732: contig of 1016 bp in length
 * 164733 164832: gap of unknown length
 * 164833 165958: contig of 1126 bp in length
 * 165959 166058: gap of unknown length
 * 166059 167610: contig of 1552 bp in length
 * 167611 167710: gap of unknown length

167711	168803:	contig of 1093 bp in length
*	168903:	gap of unknown length
*	168904:	contig of 1068 bp in length
*	169972:	170071: gap of unknown length
*	170072:	171105: contig of 1034 bp in length
*	171106:	171205: gap of unknown length
*	171206:	172289: contig of 1084 bp in length
*	172290:	172389: gap of unknown length
*	172390:	173513: contig of 1124 bp in length
*	173514:	173613: gap of unknown length
*	173614:	175296: contig of 1683 bp in length.
*		

Query Match	12.18;	Score 41.4;	DB 2;	Length 175296;	
Best Local Similarity	48.24;	Pred. No. 33;			
Matches 146;	Conservative	0;	Mismatches 156;	Indels 1;	Gaps 1;
QY 38	acctgcagacgatgggcctcagcagaacacgaggaactccagccagcagctgcgctt	97			
Db 141423	AGCAGCAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGC	141364			
QY 98	cggctggctccgacgcagcctggatcagttgtctcgccatgtttcatcatgatgatctgc	157			
Db 141363	AGCAGCAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGG	141304			
QY 158	aacagagccagggcgacgatgcaatcaggagttgtgcaacgaacacccgcagacggtc	217			
Db 141303	AGCAGGAGCAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACG	141244			
QY 218	aacaggaagccctgagtcctttgacgcagatgtgtatgcagatctgtgatcagctgatgc	277			
Db 141243	AGCAGCAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGG	141184			
QY 278	agaaccagg-gcggcgccgcataggcgcggtggcggttcgttcaacagcagcctggcgggc	336			
Db 141183	AGGACGAGGACGACGAGGACGAGGACGAGGACGAGGACGAGGACGAGGACG	141124			
QY 337	aac 339				
Db 141123	AGC 141121				

RESULT 14
 AC092203
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Jia, H., Zhang, P., Lin, S. and Roe, B. A.
 Mus musculus BAC Clone rp23-422n18
 Unpublished
 2 (bases 1 to 204152)
 Jia, H., Zhang, P., Lin, S. and Roe, B. A.
 Direct Submission
 Submitted (28-JUN-2001) Department of Chemistry And Biochemistry,
 The University of Oklahoma, 620 Parrington Oval, Room 208, Norman,
 OK 73019, USA
 On Jan 29, 2002 this sequence version replaced gi:18139426.
 ----- Genome Center
 Center: Department of Chemistry And Biochemistry
 The University Of Oklahoma
 Center code: UOKNOR

```

* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* as soon as it is available and the accession number will
* be preserved.
*
* 1
* 2023: contig of 2023 bp in length
* 2024 2123: gap of unknown length
* 2124 4788: contig of 2665 bp in length
* 4789 4889: gap of unknown length
* 4889 7174: contig of 2286 bp in length
* 7175 7274: gap of unknown length
* 7275 10351: contig of 3077 bp in length
* 10352 10451: gap of unknown length
* 10452 13314: contig of 2863 bp in length
* 13315 13414: gap of unknown length
* 13415 16003: contig of 2589 bp in length
* 16004 16103: gap of unknown length
* 16104 18443: contig of 2340 bp in length
* 18444 18543: gap of unknown length
* 18544 24743: contig of 6200 bp in length
* 24744 24843: gap of unknown length
* 24844 62943: contig of 38100 bp in length
* 62944 63043: gap of unknown length
* 63044 97892: contig of 34849 bp in length
* 97893 97992: gap of unknown length
* 97993 204152: contig of 106160 bp in length.
* Location/Qualifiers
* - 1..204152

```

Query Match	12.0%	Score 41.2;	DB 2;	Length 204152;
Best Local Similarity	49.1%;	Pred. No. 36;		
Matches 109;	Conservative 0;	Mismatches 113;	Indels 0;	Gaps 0;
Qy 64	cagcaaacacgaggactccagccagcagctcgcttcgctcgcagcagcagctgagat	123		
Db 106646	CACCAGCAGGAGCAGCAGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGG	106705		
Qy 124	cagtgctgcgcattgttcatatgatgtctgcaacagagccaggcagcgatgcaaat	183		
Db 106706	CAGCAGCAGGAGCAGCAGCAGGAGCAGCAGCAGCAGGAGCAGCAGCAGCAGC	106765		
Qy 184	caggagtgctggcacaacccagaaagcgtcaacagaaagccctgagctcgttgacg	243		
Db 106766	CAGCAGGACAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	106825		
Qy 244	cagatgctgatgcagatcgtgtatgcagctgatgcagaaccag	285		
Db 106826	CAGCAGCAGGAGCAGCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGG	106867		

RESULT	15
AC090533/c	
LOCUS	
DEFINITION	AC090533 272545 bp DNA linear HTG 25-JAN-2002 Mus musculus clone RP23-33H13 strain C57BL6/J, WORKING DRAFT SEQUENCE, 45 unordered pieces.
ACCESSION	AC090533
VERSION	AC090533.4 GI:17976439
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	house mouse.
ORGANISM	Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE	1 (bases 1 to 272545)
AUTHORS	Li, L., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M.

Goltz, J., Halder, A., Hall, L., Han, J., Ioshikhes, I.P., Lee, E.,
 Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.
 High Throughput Mouse Sequencing
 2 (bases 1 to 272545)
 Li, L., Montgomery, K.T., Grills, G., Chiu, D., Decker, J., Fusina, M.,
 Goltz, J., Halder, A., Hall, L., Han, J., Ioshikhes, I.P., Lee, E.,
 Perera, A., Shim, C., Thomas, E. and Kucherlapati, R.
 Direct Submission
 Submitted (02-MAR-2001) Department of Molecular Genetics, Albert
 Einstein College of Medicine Genome Center, 1300 Morris Park Ave.,
 Bronx, NY 10461, USA
 On Dec 21, 2001 this sequence version replaced gi:13310871.
 -----Genome Center
 Center: Harvard Partners Genome Center
 Center Code: HPGC
 Web site:
 http://wchanning.bwh.harvard.edu:9088/hpcgg/jsp/hpcgg/Sequence/mous
 e.html
 Contact: gntm@capecod.bwh.harvard.edu
 -----Summary Statistics

Center project name: ADU
 Sequencing vector: pUC18; L08752
 Chemistry: Dye-terminator Big Dye; 100%
 *Consensus quality: 257441 at least Q20
 *Consensus quality: 253718 at least Q30
 *Consensus quality: 247095 at least Q40
 Estimated insert size: agarose-FP - N/A
 **Estimated insert size: 271665 - sum-of-contigs
 Quality coverage: agarose-FP - N/A
 Quality coverage: 6.8 x in Q20 bases; sum-of-contigs estimation

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 45 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1 30267: contig of 30267 bp in length
 * 30268: gap of unknown length
 * 30288: contig of 28113 bp in length
 * 58400: gap of unknown length
 * 58420: gap of unknown length
 * 58421: contig of 30444 bp in length
 * 88865: gap of unknown length
 * 88885: contig of 21005 bp in length
 * 109890: gap of unknown length
 * 109910: contig of 15224 bp in length
 * 125133: gap of unknown length
 * 125154: contig of 14396 bp in length
 * 139550: gap of unknown length
 * 139570: contig of 13841 bp in length
 * 153410: gap of unknown length
 * 153431: contig of 9137 bp in length
 * 162568: gap of unknown length
 * 162588: contig of 8457 bp in length
 * 171045: gap of unknown length
 * 171065: contig of 8240 bp in length
 * 179305: gap of unknown length
 * 179325: contig of 7735 bp in length
 * 187059: gap of unknown length
 * 187060: contig of 10519 bp in length
 * 187080: gap of unknown length
 * 197599: contig of 8348 bp in length
 * 197619: gap of unknown length
 * 205967: gap of unknown length
 * 205987: contig of 6463 bp in length
 * 212450: gap of unknown length
 * 212470: contig of 5283 bp in length
 * 217753: gap of unknown length
 * 217773: contig of 6171 bp in length
 * 223944: gap of unknown length
 * 223963: contig of 6399 bp in length
 * 230362: contig of 5869 bp in length

230363: gap of unknown length
 * 230383: contig of 5869 bp in length
 * 236252: gap of unknown length
 * 236271: contig of 4301 bp in length
 * 240572: gap of unknown length
 * 240573: contig of 2197 bp in length
 * 240593: gap of unknown length
 * 242789: contig of 1886 bp in length
 * 242809: gap of unknown length
 * 242810: contig of 1886 bp in length
 * 244695: gap of unknown length
 * 244696: contig of 2254 bp in length
 * 244716: gap of unknown length
 * 246969: contig of 881 bp in length
 * 246970: gap of unknown length
 * 247871: contig of 303 bp in length
 * 247890: gap of unknown length
 * 247891: contig of 303 bp in length
 * 248194: gap of unknown length
 * 248213: contig of 1909 bp in length
 * 250123: gap of unknown length
 * 250124: contig of 1225 bp in length
 * 250143: gap of unknown length
 * 251368: contig of 1362 bp in length
 * 251388: gap of unknown length
 * 252749: contig of 871 bp in length
 * 252750: gap of unknown length
 * 252770: contig of 871 bp in length
 * 253641: gap of unknown length
 * 253660: contig of 1372 bp in length
 * 253661: gap of unknown length
 * 255032: contig of 2281 bp in length
 * 255033: gap of unknown length
 * 255053: contig of 2056 bp in length
 * 257334: gap of unknown length
 * 257354: contig of 1030 bp in length
 * 259410: gap of unknown length
 * 259429: contig of 1030 bp in length
 * 259430: gap of unknown length
 * 260460: contig of 350 bp in length
 * 260479: gap of unknown length
 * 260480: contig of 812 bp in length
 * 260830: gap of unknown length
 * 260849: contig of 812 bp in length
 * 261661: gap of unknown length
 * 261682: contig of 851 bp in length
 * 262532: gap of unknown length
 * 262533: contig of 799 bp in length
 * 262553: gap of unknown length
 * 263352: contig of 846 bp in length
 * 263371: gap of unknown length
 * 263372: contig of 968 bp in length
 * 264217: gap of unknown length
 * 264237: contig of 968 bp in length
 * 265205: gap of unknown length
 * 265206: contig of 1045 bp in length
 * 265226: gap of unknown length
 * 266271: contig of 459 bp in length
 * 266290: gap of unknown length
 * 266750: gap of unknown length
 * 266759: contig of 949 bp in length
 * 266770: gap of unknown length
 * 267719: contig of 838 bp in length
 * 267739: gap of unknown length
 * 268577: gap of unknown length
 * 268596: contig of 1414 bp in length
 * 268597: gap of unknown length
 * 270011: contig of 1624 bp in length
 * 270031: gap of unknown length
 * 271655: contig of 871 bp in length
 * 271675: gap of unknown length
 * 272545: contig of 871 bp in length.
 FEATURES
 source
 Location/Qualifiers
 1. 272545
 /organism="Mus musculus"
 /strain="C57BL6/J"
 /db_xref="taxon:10090"
 /clone="RP23-33H13"
 /sex="male"
 misc_feature
 1. 30267
 /note="assembly_name:Contig108"
 30288..58400
 /note="assembly_name:Contig107"
 clone_end:SP6
 vector_side:left
 58421..88864
 /note="assembly_name:Contig106"
 88885..109889
 /note="assembly_name:Contig105"

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 10:22:48 ; Search time 177.26 Seconds
(without alignments)
3312.562 Million cell updates/sec

Title: US-09-829-124-1

Perfect score: 342

Sequence: 1 atggactctatcggaacaa.....gcagctggcggaacgcc 342

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_032802.*
1: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS1/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	342	100.0	342	22	AAD18127
2	38.8	11.3	2004	18	AA185356
3	38.4	11.2	1925	20	AA190924
4	38.4	11.2	3840	23	ABL23322
5	38	11.1	3121	21	AA151260
6	37.8	11.1	1581	22	AA161055
7	37.8	11.1	4335	23	AA179854
8	37.8	11.1	5120	22	AA184677
9	37.6	11.0	1035	18	AA149316

10	37.6	11.0	1035	19	AA154609
11	37.6	11.0	1035	19	AA139975
12	37.6	11.0	1035	19	AA136430
13	37.6	11.0	1035	20	AA183991
14	37.6	11.0	1035	21	AA100674
15	37.6	11.0	1035	21	AA114944
16	37.6	11.0	1035	22	AA12808
17	37.4	10.9	1650	23	ABL23323
18	37.4	10.9	2444	23	ABL05117
19	37.4	10.9	4444	23	ABL05116
20	37.2	10.9	447	23	ABL27665
21	37.2	10.9	506	21	AA12475
22	37.2	10.9	574	22	AA101496
23	37.2	10.9	2338	12	AA14183
24	37.2	10.9	2338	19	AA123249
25	37.2	10.9	2338	21	AA138195
26	37.2	10.9	2744	16	AA108470
27	37.2	10.9	3477	23	ABL27664
28	37.2	10.9	4203	23	AA105260
29	37	10.8	900	23	AA151554
30	37	10.8	6463	21	AA145239
31	36.8	10.8	477	22	AA101459
32	36.6	10.7	2177	21	AA157853
33	36.6	10.7	6155	23	AA109166
34	36.6	10.7	6955	22	AA172397
35	36.6	10.7	13010	22	AA172395
36	36.4	10.6	317	21	AA102116
37	36.4	10.6	821	22	AA105110
38	36.4	10.6	4105	22	AA161196
39	36.4	10.6	4396	22	AA17588
40	36.4	10.6	4428	22	AA159410
41	36.4	10.6	4403765	22	AA199683
42	36.2	10.6	396	22	AA181388
43	36.2	10.6	423	22	AA164895
44	36	10.5	1126	19	AA129159
45	35.8	10.5	4262	21	AA164577

ALIGNMENTS

RESULT 1

AAD18127
ID AAD18127 standard; cDNA; 342 BP.

XX AC AAD18127;

XX DT 18-DEC-2001 (first entry)

XX DE Xanthomonas campestris hypersensitive response elicitor protein cDNA.

XX KW Plant pathogen; hypersensitive response elicitor; HRE; stress resistance; disease resistance; Erwinia amylovora hypersensitive response elicitor; harpin; signal transduction; growth enhancement; insect control; virucide; fungicide; antibacterial; ss.

XX OS Xanthomonas campestris.

XX FH key Location/Qualifiers

XX FT CDS 1..342

XX FT /*tag= a

XX FT /product= "Xanthomonas campestris hypersensitive response elicitor protein"

XX FT /note= "CDS does not include stop codon"

XX FT /partial

XX PN WO200170988-A2.

XX XX 27-SEP-2001.

XX XX 19-MAR-2001; 2001WO-US08728.

XX XX 23-MAR-2000; 2000US-191649P.

Pseudomonas solana
Pseudomonas solana
Hypersensitive res
DNA encoding a hyp
Pseudomonas solana
DNA encoding a hyp
Pseudomonas solana
Pseudomonas solana
Drosophila melanog
Drosophila melanog
Drosophila melanog
Human secreted exp
Human reproductive
N. clavipes draglin
Nephila clavipes s
N. clavipes spider
MISPI-containing p
Drosophila melanog
Human reproductive
Pseudomonas aerugi
DNA encoding angio
Human reproductive
Protein regulating
Drosophila melanog
Human immune/haema
Human immune/haema
Human secreted pro
Human cDNA clone (
Human polynucleoti
Human cDNA sequenc
Human polynucleoti
Mycobacterium tube
Xanthomonas hrif p
Xanthomonas hrif p
Nucleotide sequenc
E. tenella CGMP de

01-DEC-2000; 2000US-250710P.
(EDEN-) EDEN BIOSCIENCE CORP.
Song X, Fan H, Wei Z;
WPT; 2001-5901177/66.
P-PSDB; AAEE10807.
New plant pathogen hypersensitive response elicitor-receptor protein isolated from plants, which upon silencing is used to study plant signal transduction pathways leading to disease resistance and growth enhancement
Disclosure; Page 19; 78pp; English.
The present invention relates to an isolated protein which serves as a receptor in plants for plant pathogen hypersensitive response elicitor (HRE). The plants are made resistant to infection by viruses, bacteria and fungi and are imparted with resistance against environmental stress and insects through HRE treatment. The protein is useful for understanding the harpin (Erwinia amylovora hypersensitive response elicitor) induced signal transduction pathway in plants. The protein is useful for studying the downstream components of signal transduction pathway in plants which eventually leads to disease resistance, growth enhancement, insect control and stress resistance. Imparting disease resistance to plants through HRE treatment has the potential to treat previously untreatable diseases, treating diseases systemically which might not be treated separately due to cost and avoids the use of infectious agents or environmentally harmful materials. By HRE treatment enhanced plant growth is achieved which includes greater yield, increased quantity of seeds produced, increased percentage of seeds germinated, increased plant size, greater biomass, more and bigger fruits, etc. which results in economic benefit to cultivators. Greater yield, increased size and enhanced biomass allow greater revenue generation from the given plot of plant. The present sequence is *Xanthomonas campestris* pv. *pelargonii* hypersensitive response elicitor protein cDNA.
Sequence 342 BP; 83 A; 94 C; 109 G; 56 T; 0 other;

AA85356 standard; DNA; 2004 BP.	
AA85356;	
09-DEC-1997 (first entry)	
Nephila clavipes spider silk protein 2 Kb DNA sequence.	
High strength film; fibre; woven article; parachutes; sails; absorber; body armour; heavy metal; biological weapon; chemical; flavour; fragrance; Nephila clavipes; ss.	
Nephila clavipes.	
Key Location/Qualifiers	
CDS 40..1980	
FT: /*tag= a	
FT: /product= Silk_protein	
WO9708315-A1.	
06-MAR-1997.	
22-AUG-1996; 96WO-US13767.	
22-AUG-1995; 95US-0517694.	
(BASE/) BASEL R M.	
(ELIO/) ELION G R.	
Basel RM, Elion GR;	
WPI; 1997-179272/16.	
P-PSDB; AAW2178.	
New opt. multimerised DNA sequences encoding spider silk protein - contg. both repetitive and non-repetitive sequences, useful for making high strength films, fibres, woven articles etc.	
Claim 14; Fig 1; 57pp; English.	
A process has been developed for the production of a DNA fragment encoding silk protein. The process involves: (a) selecting target DNA, from a silk-producing spider, that contains many repetitive and non-repetitive regions; (b) selecting a single-stranded DNA primer of at least 10 nucleotides with a sequence that is complementary to a region of the target; (c) repetitively combining the primer with melted target DNA, incubating the mixture with nucleotides and a DNA polymerase with proofreading activity to produce a DNA fragment which is complementary to the target and is at least 2 kb long. The present sequence represents a 2 kb DNA sequence which encodes the spider silk protein from Nephila clavipes. The DNA fragment can be used to make fibres, films, woven articles, e.g. for use in parachutes, sails, body armour, and absorbers (e.g. of heavy metals, biological weapons, DNA, chemicals, flavours and fragrances). The high molecular weight (90-250 kD) of spider silk proteins can be produced on a commercial scale (at over 2 g/l cell mass). It has better tensile strength and elasticity than silkworm silk. Inclusion of both repetitive and non-repetitive regions ensures isolation of stable clones.	
Sequence 2004 BP; 481 A; 386 C; 791 G; 346 T; 0 other:	

D	b	2165	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagca	2224
Q	y	228	cctgagtccgttgacgcagatgctgatgcagatcgtgatgcagctgatagcagaaccagg-	286
D	b	2225	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagca	2284
Q	y	287	gcggcgccgcgaTggcggtggcggttcggtcaacacagcagcctgggcggcaac	339
D	b	2285	gcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagcagc	2337
R	E	S	U	L
R	E	S	U	L
I	D	AAT49316	standard; DNA; 1035 BP.	
A	C	AAT49316;		
D	T	30-MAR-1997	(first entry)	
D	E	Hypersensitive response elicitor gene.		
K	W	Hypersensitive response; elicitor; Pseudomonas solanacearum; plant;		
K	W	disease-resistance; Escherichia coli; infiltration; virus;		
K	W	bacterium; fungus; pathogen; biological control agent; ss.		
O	S	Pseudomonas solanacearum.		
P	N	WO9639802-A1.		
P	D	19-DEC-1996.		
P	F	05-JUN-1996; 96WO-US08819.		
P	R	07-JUN-1995; 95US-0475775.		
P	A	(CORR) CORNELL RES FOUND INC.		
P	I	Beer SV, Wei Z;		
D	R	WPI; 1997-051614/05.		
D	R	P-PSDB; AAW06600.		
P	T	Imparting pathogen resistance to plants - with hypersensitive		
P	T	response elicitor polypeptide or protein		
X	S	Disclosure; Page 51-52; 69pp; English.		
X	S	This sequence encodes a hypersensitive response elicitor from		
C	C	Pseudomonas solanacearum. The elicitor may be used in a new method		
C	C	for imparting pathogen resistance to plants, by application of the		
C	C	elicitor in a non-infectious form to plant cells, by spraying,		
C	C	injection, leaf abrasion, or plant infection with recombinant		
C	C	bacteria (non-infectious to the host plant, e.g. Escherichia coli)		
C	C	expressing the elicitor as a biological control agent, to allow		
C	C	recombinant protein infiltration into the plant. The method		
C	C	confers virus, bacterium or fungus disease-resistance on crops and		
C	C	ornamental plants.		
S	Q	Sequence 1035 BP; 227 A; 330 C; 368 G; 110 T; 0 other;		

SQ	Sequence	1035 BP;	227 A;	330 C;	368 G;	110 T;	0 other;		
Query Match	11.0%;	Score	37.6;	DB	18;	Length	1035;		
Best Local Similarity	55.3%;	Pred.	No.	2.7;	Mismatches	0;	Gaps	0;	
Matches	73;	Conservative	0;	Mismatches	59;	Indels	0;	Gaps	0;
QY	211	aacggtcaacagaaaggccttgagtccgttacgcagcatgctgatcgatcgatgtagcg	270						
Db	712	agcgagaaccaggcgccctcacccggcggtgctgaaggctgatgaagattctgaacgcg	771						
QY	271	ctgatgcagaaaccaggcgccggcgatggcggttggcggtttcgggtcaacagcagcctg	330						
Db	772	ctggtgcagatgatgcagcaaggcgacctggcggcgccaaccaggcgcgagggcggtcg	831						

QY 331 ggcggcaacgcc 342
 |||
 Db 832 aaggggtgccggc 843

RESULT 10
 AAV54609
 ID AAV54609 standard; DNA; 1035 BP.

XX AC AAV54609;

XX DT 07-DEC-1998 (first entry)

XX DE Pseudomonas solanacearum hypersensitive response elicitor DNA.

XX KW Hypersensitive response elicitor; HRE; insect resistance;
 KW biological control; transgenic plant; ss.

XX OS Pseudomonas solanacearum.

XX PN WO9837752-A1.

XX PD 03-SEP-1998.

XX PF 26-FEB-1998; 98WO-US03604.

XX PR 28-FEB-1997; 97US-0039226.

XX PA (CORR) CORNELL RES FOUND INC.

XX PI Wei Z, Zitter TA;

XX DR WPI: 1998-495374/42.
 DR P-PSDB; AAW75865.

XX PT Use of hypersensitive response elicitor polypeptide - for
 PT application to plants or seeds or transgenic plants or seeds for the
 PT control of insects.

XX PS Disclosure; Page 15; 75pp; English.

XX CC This DNA sequence comprises an open reading frame that encodes a
 CC hypersensitive response elicitor (HRE) (see AAW75865 of Pseudomonas
 CC solanacearum. The invention relates to the use of a HRE
 CC polypeptide or protein to control insects on plants or plants grown
 CC from seed treated with HRE. Also claimed is a method of insect
 CC control for plants that involves: (a) providing a transgenic plant
 CC or seed transformed with a DNA molecule (see AAV54606-09) encoding a
 CC HRE polypeptide or protein (see AAW75862-67); and (b) growing the
 CC transgenic plants or transgenic plants produced from the transgenic
 CC seeds to control insects. HRE prevents direct insect damage to
 CC plants by feeding injury. It kills insects close to plants, and
 CC interferes with insect larval feeding on such plants. It also
 CC prevents insects from colonising host plants and releasing
 CC phytotoxins which result in disease damage to plants.

XX SQ Sequence 1035 BP; 227 A; 330 C; 368 G; 110 T; 0 other;

Query Match 11.0%; Score 37.6; DB 19; Length 1035;
 Best Local Similarity 55.3%; Pred. No. 2.7;
 Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 211 aacggtcaacaggaagcctgagtcggttgacgcagatgctgacagatgctgagcg 270

Db 712 agcgaagaccaggcgccctaccggtgctgcaaaagctgatgaagatcctgaacgcg 771

QY 271 ctgatgcgaacaccaggcgccgctgagcggtggtggttcgtcaacagcagcctg 330

Db 772 ctggtgcagatgatgcagcaagcgccctcgcgcgccgcaacagcgcgagggcggtcg 831

QY 331 ggcggcaacgcc 342
 |||

Db 832 aaggggtgccggc 843

RESULT 11

AAV39975

XX ID AAV39975 standard; DNA; 1035 BP.

XX AC AAV39975;

XX DT 09-NOV-1998 (first entry)

XX DE Pseudomonas solanacearum hypersensitive response elicitor DNA.

XX KW Hypersensitive response elicitor; HRE; growth; transgenic plant;
 KW ss.

XX OS Pseudomonas solanacearum.

XX PN WO9832844-A1.

XX PD 30-JUL-1998.

XX PF 27-JAN-1998; 98WO-US01507.

XX PR 27-JAN-1997; 97US-0036048.

XX PA (CORR) CORNELL RES FOUND INC.

XX PI Beer SV, Qiu D, Wei Z;

XX DR WPI: 1998-427940/36.
 DR P-PSDB; AAW62457.

XX PT Method for enhancing plant growth - comprises use of hypersensitive
 PT response elicitor polypeptide or protein which may also effect, e.g.
 PT increase in plant height or earlier germination seed
 XX PS Disclosure; Page 21; 110pp; English.

XX CC This DNA sequence comprises an open reading frame that encodes a
 CC hypersensitive response elicitor (HRE) (see AAW62457) of Pseudomonas
 CC solanacearum. A method of enhancing growth in plants comprises:
 CC (a) applying a HRE polypeptide or protein in a non-infectious form
 CC to a plant or plant seed under conditions effective to enhance
 CC growth of the plant or plants grown from the seed, or (b) providing
 CC a transgenic plant or plant seed transformed with a DNA molecule
 CC encoding a HRE polypeptide or protein, and growing the transgenic
 CC plant or a plant produced from the transgenic seed under conditions
 CC effective to enhance plant growth. HRES (see AAW62454-59) or nucleic
 CC acids (see AAV62454-57) encoding them can be used to increase plant
 CC growth. The HRES may also result in increased plant height and
 CC yield, and effect early germination and maturation of plant seed
 CC and early colouration of fruit and plants.

XX SQ Sequence 1035 BP; 227 A; 330 C; 368 G; 110 T; 0 other;

Query Match 11.0%; Score 37.6; DB 19; Length 1035;
 Best Local Similarity 55.3%; Pred. No. 2.7;
 Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 211 aacggtcaacaggaagcctgagtcggttgacgcagatgctgacagatgctgagcg 270

Db 712 agcgaagaccaggcgccctaccggtgctgcaaaagctgatgaagatcctgaacgcg 771

QY 271 ctgatgcgaacaccaggcgccgctgagcggtggtggttcgtcaacagcagcctg 330

Db 772 ctggtgcagatgatgcagcaagcgccctcgcgcgccgcaacagcgcgagggcggtcg 831

QY 331 ggcggcaacgcc 342
 |||

Db 832 aaggggtgccggc 843

RESULT	12
AAV36430	
AD AAV36430 standard; DNA; 1035 BP.	
XX AC	AAV36430;
XX DT	26-OCT-1998 (first entry)
XX DE	Hypersensitive response elicitor
XX KW	Hypersensitive response elicitor
XX KW	pathogen resistance; disease res
XX OS	Pseudomonas syringae.
XX PN	W09824297-A1.
XX PD	11-JUN-1998.
XX PF	04-DEC-1997; 97WO-US22629..
XX PR	05-DEC-1996; 96US-0033230.
XX PA	(CORR) CORNELL RES FOUND INC.
XX PI	Beer SV, Qiu D, Wei Z;
XX DR	WPI; 1998-332931/29.
XX DR	P-PSDB; AA61116.
XX PT	Imparting pathogen resistance to
XX PT	hypersensitive response elicitor
XX PS	Disclosure; Page 24-25; 85pp; En
XX CC	This nucleotide sequence from PstI
XX CC	a coding region for a hypersensit
XX CC	AA61115). The invention relates
XX CC	hypersensitive response induced
XX CC	of seeds. Isolated HRE polypept
XX CC	means of imparting pathogen resi
XX CC	seeds. Alternatively, bacteria
XX CC	HRE can be applied to the plant s
XX CC	containing a DNA molecule encodi
XX CC	are used. HRE polypeptide sequen
XX CC	Erwinia amylovora, Pseudomonas s
XX CC	Xanthomonas campestris pv. glycini
XX CC	pelargonii (see AA61113-18) are
XX CC	pathogen resistance without using
XX CC	environment or pathogenic to the
XX CC	adjacent plants.
XX SQ	Sequence 1035 BP; 227 A; 330 C; 3

[illegible]

RESULT 13	
AAV83991	
ID	AAV83991 standard; DNA; 1035 BP.
XX	
AC	AAV83991;
XX	
DT	09-WAR-1999 (first entry)
XX	
DE	DNA encoding a hypersensitive response elicitor protein.
XX	
KW	Hypersensitive response elicitor protein; hairpin protein;
KW	disease resistance; seed quality; insect control; corn borer;
KW	Lepidoptera larvae; transgenic plant; ss.
XX	
OS	Pseudomonas solanacearum.
XX	
PH	Key
XX	Location/Qualifiers
FT	1..1035
FT	/*tag a
FT	/product= Hypersensitive_response_elicitor_protein
XX	
PN	WO9854214-A2.
XX	
PD	03-DEC-1998.
XX	
PF	28-MAY-1998; 98WO-US10874.
XX	
PR	30-MAY-1997; 97US-0048109.
XX	
PA	(CORR) CORNELL RES FOUND INC.
PA	(EDEN-) EDEN BIOSCIENCE CORP..
XX	
PI	Beer SV, Laby RJ, Wei Z;
XX	
WPI	WPI; 1999-070210/06.
DR	P-PSDB; AAW87641.
XX	
PT	New fragments of an Erwinia hypersensitive response elicitor protein
PT	and related DNA - used to impart disease resistance to plants, to
PT	increase their growth and to control insects
XX	
PS	Disclosure; Page 16-17: 94pp; English.
XX	
CC	The present sequence encodes a hypersensitive response elicitor
CC	protein (also called hairpin protein) that is able to elicit a
CC	hypersensitive response in plants. The specification also describes
CC	hypersensitive response elicitors from other pathogenic organisms.
CC	The protein, in non-infectious form, is applied to plants to impart
CC	disease resistance (to a wide range of viral, bacterial and fungal
CC	pathogens), to improve growth (yield, quantity and quality of seed
CC	to provide earlier germination etc.) and to control insects (e.g. c
CC	borers, Lepidoptera larvae etc.) The same results are provided by
CC	transgenic plants expressing the protein.
XX	
SQ	Sequence 1035 BP; 227 A; 330 C; 368 G; 110 T; 0 other.

	Query Match	11.0%	Score 37.6;	DB 20;	Length 1035;
	Best Local Similarity	55.3%;	Pred. NO. 2.7;	Mismatches 0;	Gaps 0;
	Matches 73;	Conservative 0;	Mismatches 59;	Indels 0;	Gaps 0;
Qy	211	aacgggtcaaacaggagcgctgagtccgttgacgcagatcgtgatgcagatcgtgatgcag	270		
Dd	712	agcgaagaccaggcgacctaccgcggtgctgcacaagactgatgaagatcctgaacgcg	771		
Qy	271	ctgatgcagaaccacagggccgcgcggcatggcggttgcggtttcggccaacagcacctg	330		
Dd	772	ctggtgcagatgatgcagcaaggcgccctcgttcggcggaaccaaccaggcgagggcggtctg	831		
Qy	331	ggcgggcaacgcc	342		
Dd	832	aagggtgccgcg	843		

Search completed: June 28, 2002, 11:25:04
Job time: 3736 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 10:22:23 ; Search time 39.55. Seconds.
(without alignments)
2124.062 Million cell updates/sec

Title: US-09-829-124-1

Perfect score: 342

Sequence: 1 atggactctatcggaacaa.....gcagcctggcggaagcc.342

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 303533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/PTUS_COMB1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	37.6	11.0	1035	1	US-08-891-254-8	Sequence 8, Appli
2	37.6	11.0	1035	2	US-08-819-539-8	Sequence 8, Appli
3	37.6	11.0	1035	2	US-09-030-270A-8	Sequence 8, Appli
4	37.6	11.0	1035	4	US-08-984-207-8	Sequence 8, Appli
5	37.6	11.0	1035	4	US-09-013-587-8	Sequence 8, Appli
6	37.6	11.0	1035	5	PCT-US96-08819-8	Sequence 8, Appli
7	37.2	10.9	2338	1	US-08-425-069-1	Sequence 1, Appli
8	37.2	10.9	2338	2	US-08-317-844B-1	Sequence 1, Appli
9	36.4	10.6	4403765	4	US-09-103-840A-2	Sequence 2, Appli
10	36	10.5	657	4	US-08-998-416-1132	Sequence 1132, Ap
11	36	10.5	1126	4	US-08-949-155-5	Sequence 5, Appli
12	35.2	10.3	1665	3	US-08-881-784-8	Sequence 8, Appli
13	35.2	10.3	1665	4	US-09-292-768-3	Sequence 3, Appli
14	35.2	10.3	1665	4	US-09-292-768-3	Sequence 3, Appli
15	35.2	10.3	1665	4	US-09-292-768-67	Sequence 67, Appli
16	34.6	10.1	712	1	US-09-292-768-69	Sequence 69, Appli
17	34.6	10.1	712	1	US-08-276-452A-72	Sequence 72, Appli
18	34	9.9	543	6	US-08-798-744-72	Sequence 72, Appli
19	33.8	9.9	1995	1	5273901-6	Patent No. 5273901
20	33.8	9.9	1995	2	US-08-425-069-3	Sequence 3, Appli
21	33.4	9.8	1995	2	US-08-317-844B-3	Sequence 3, Appli
22	33.2	9.7	397	3	US-09-056-556-221	Sequence 221, App
23	33.2	9.7	397	3	US-09-253-691-3	Sequence 3, Appli
24	33.2	9.7	3044	1	US-07-872-644-44	Sequence 44, Appli
25	33.2	9.7	3044	1	US-08-297-494-44	Sequence 44, Appli
26	33.2	9.7	3044	1	US-08-297-510-44	Sequence 44, Appli
27	33.2	9.7	3044	1	US-08-479-532-44	Sequence 44, Appli
28	33.2	9.7	3044	1	US-08-455-526-44	Sequence 44, Appli

C	28	33.2	9.7	3044	1	US-08-455-525-44	Sequence 44, Appli
C	29	33.2	9.7	3044	3	US-09-139-491-44	Sequence 44, Appli
C	30	33.2	9.7	3044	5	PCT-US92-03222-44	Sequence 44, Appli
	31	33.2	9.7	3468	1	US-07-951-715A-2	Sequence 2, Appli
	32	33.2	9.7	3468	1	US-07-951-715A-4	Sequence 4, Appli
	33	33.2	9.7	3468	2	US-08-459-448A-2	Sequence 2, Appli
	34	33.2	9.7	3468	2	US-08-459-448A-4	Sequence 4, Appli
	35	33.2	9.7	3468	3	US-08-459-595A-2	Sequence 4, Appli
	36	33.2	9.7	3468	3	US-08-459-595A-4	Sequence 4, Appli
	37	33.2	9.7	3468	3	US-08-459-504B-2	Sequence 2, Appli
	38	33.2	9.7	3468	3	US-08-459-504B-4	Sequence 4, Appli
	39	33.2	9.7	3468	3	US-08-459-444-2	Sequence 2, Appli
	40	33.2	9.7	3468	3	US-08-459-444-4	Sequence 4, Appli
	41	33.2	9.7	3468	3	US-09-053-549-3	Sequence 3, Appli
	42	33.2	9.7	3468	3	US-09-053-549-5	Sequence 5, Appli
	43	33.2	9.7	3468	4	US-09-547-422-2	Sequence 2, Appli
	44	33.2	9.7	3468	4	US-09-547-422-4	Sequence 4, Appli
	45	33	9.6	1515	4	US-09-292-768-5	Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-891-254-8
: Sequence 8, Application US/08891254
: Patent No. 5776889
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: APPLICANT: Beer, Steven V.
: TITLE OF INVENTION: Hypersensitive Response
: TITLE OF INVENTION: Induced Resistance In Plants
: NUMBER OF SEQUENCES: 9
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Nixon, Hargrave, Devans & Doyle
: STREET: Clinton Square, P.O. Box 1051
: CITY: Rochester
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 14603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/891,254
: FILING DATE: 10-JUL-1997
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/475,775
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldman, Michael L.
: REGISTRATION NUMBER: 30,727
: REFERENCE/DOCKET NUMBER: 14603/10050
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (716) 263-1304
: TELEFAX: (716) 263-1600
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1035 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
US-08-891-254-8

Query Match 11.0%; Score 37.6; DB 1; Length 1035;
Best Local Similarity 55.3%; Pred. No. 0.19; Indels 0; Gaps 0;
Matches 73; Conservative 0; Mismatches 59;

RESULT 5
US-09-013-587-8
; Sequence 8, Application US/09013587
; Patent No. 6277814
; GENERAL INFORMATION:
; APPLICANT: Qiu, Dewen
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: ENHANCEMENT OF GROWTH IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans' & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.

```

RESULT      6
PCT-US96-08819-8
; Sequence 8, Application PC/TUS9608819
; GENERAL INFORMATION:
; APPLICANT: Cornell Research Foundation, Inc.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
;                     RESISTANCE IN PLANTS
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08819
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,775

```

```

RESULT      8
US-08-317-844B-1
: Sequence 1, Application US/08317844B
: Patent No. 5969894
: . GENERAL INFORMATION:
:   APPLICANT: Lewis, Randolph V.
:   APPLICANT: Xu, Ming
:   APPLICANT: Hinman, Michael B.
:   TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
:   TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
:   TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
:   NUMBER OF SEQUENCES: 62
:   CORRESPONDENCE ADDRESS:
:   ADDRESSER: Birch, Stewart, Kolasch & Birch
:   STREET: 301 No. 5989894th Washington Street
:   CITY: Falls Church
:   STATE: Virginia
:   COUNTRY: U.S.A.
:   ZIP: 22046
:   COMPUTER READABLE FORM:
:   MEDIUM TYPE: Floppy disk
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: PatentIn Release #1.0, Version #1.25
:   CURRENT APPLICATION DATA:
:   APPLICATION NUMBER: US/08/317-844B

```

```

; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" base
; OTHER INFORMATION: represented
US-09-103-840A-2

```

	Query Match	10.6%	Score 36.4	DB 4	Length 4403765
	Best Local Similarity	45.2%	Pred. No. 6.1		
	Matches 133	Conservative	0	Mismatches 161	Indels 0
	Gaps	0			
Qy	32	tcggcaactgtcagacagtgggcatcggcctcagcaacacaggagactccagccagcagt	91		
Db	2543256	TTGGCTCATCCCAANTGCGAGCGCGGACGCCGCCGCCGCCAGCCAGCGAGGCCACCCCGGAGGTAG	2543197		

QY 92 cgctctcggtgctcgcagcagcagctggtatcagttgtctgcgcatttcatcatgatga 151
 Db 2543196 CCGCGTGTGGGTGACGGGGCAATGTCAGGTGCGCGCATGSCGCGATCCGCCCTGCTGC 2543137
 QY 152 tgcgtcaacagagccaggcagcgatgcaaatcaggagtggtggcaacgacaaaccgcaga 211
 Db 2543136 CAGCGCAGCAGACCCATTCCGCTGTGTAGAGATTGCCAAATCGCGAGCACCCACCGAGC 2543077
 QY 212 acgggtcaacaggagccctgagtcctgttgacgcagatgctgtgatgcaagatcgtgatgcgc 271

QY 272 TgATGAGAAcCaggGcgGCGccgcgcTgtggcggttcggtcaacagca 325
||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 2543016 GCCAGCAGGCCTAGAGCGGTGGCGTGCCAGGCCAGGAAGTCCTGTTGGCA 2542963

```

RESULT 10
US-08-998-416-1132
; Sequence 1132, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSYPII
; TITLE OF INVENTION: AND USES THEREOF
; NUMBER OF SEQUENCES: 1152

```

COUNTRY: USA
 ZIP: 27709
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/998.416
 FILING DATE: 24-DEC-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: CH 0016/97
 FILING DATE: 31-DEC-1996
 ATTORNEY/AGENT INFORMATION:

Query Match 10.5%; Score 36; DB 4; Length 1126;
Best Local Similarity 49.0%; Pred. No. 0.54;
Matches 96; Conservative 0; Mismatches 100; Indels

OY	106	tccgaagcagcagctgatacgttgtctgcgcatttcatcatgatgctgcaacagagc	165
Db	231	TCTGACCAAGTGTGCAGGAGGAGTGCTTCAGCACCAAGTCCACCAGAACTGACGGAGCTG	290
OY	166	caggcgagcgaatgcgaatcagaagtggtggccaacaaaccacgaacaggtcaacaggaa	225
Db	291	ATAGAGCAGCATGAAGGAGGTGAAGGCCCTACCGGAGGAGCTGGAGCGCAGTGGCG	350
OY	226	ggcctgagtccgttgacgcagatgctgatgcagatcgtgatcagctgatgcagaaccag	285
Db	351	CCCGTGAACCCAGGAGACGCGAGCGGCCCTGTCCAAGGAGCTGCAGCGCGGCGAGCCCGC	410
OY	286	ggcggcgccggcattgg	301
Db	411	GTTGGCGCGACATTG	426

RESULT 12
US-08-881-784-8
: Sequence 8, Application US/08881784
: Patent No. 6083731
: GENERAL INFORMATION:
: APPLICANT: Croteau, Rodney B.
: APPLICANT: Lupien, Shari L.
: APPLICANT: Karp, Frank
: TITLE OF INVENTION: RECOMBINANT MATERIALS AND METHODS FOR
: THE PRODUCTION OF LIMONENE HYDROXYLASES
: NUMBER OF SEQUENCES: 58
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Christensen, O'Connor, Johnson and Kindness
: ADDRESSEE: PLLC
: STREET: 1420 Fifth Avenue, Suite 2800
: CITY: Seattle
: STATE: WA
: COUNTRY: USA
: ZIP: 98101
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/881,784
: FILING DATE:
: CLASSIFICATION: 435
: ATTORNEY/AGENT INFORMATION:
: NAME: Shelton, Dennis K.
: REGISTRATION NUMBER: 26,997
: REFERENCE/DOCKET NUMBER: WSUR19777
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (206) 224-0718
: TELEFAX: (206) 224-0779
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1665 base pairs
: TYPE: nucleic acid

55	atcgggctcagcaacacagagactccagccagctgccttcggctcgcagcag	114
QY		
56	atcgggctcagcaacacagagactccagccagctgccttcggctcgcagcag	115
DB		
85	atcagcaatggcgaaaccgaaaccccaagagaacctgcctccggccgcgaagctg	144
DB		
115	cagctggatcagttcgtcgccattgttcatactgatagtctcaacagagccagggcagc	174
QY		
145	cgcgtgatcggcaacctcacctcctcctatagggaagagctgcgcagcaacgctgcccacg	204
DB		

Qy	55	atcggcgctcagcaacacagaggactccagccagcagctgcctctcgctggtcctccgagcag	114
Db	85	atcaagcaattgggaaacccaacccaagaaacactgctccgggcccgcgaagctg	144
Qy	115	cagctggatcagttgctcgccatttcacatcatgatactcaacagcagcaggcagc	174
Db	145	ccgctgatcgggcacctccacctcctctatgggggaagctccgcgacacgcctggccagc	204
Qy	175	gatgcaaatcagagtgctggcaacgaacccgacgaacggtcaacaggaaagcctgagt	234
Db	205	gtggcgaacagtcacggcccgactggcgccacgtgcagctcggcgaggtgtctccgctg	264
Qy	235	ccggttacgcagatgctgatgcagatcgttgatgcagctaatcagaacacagagcggcgc	294
Db	265	ctctctccgcgagggccacgaaggagtcgatgaagctggtggaccggcctgcgcgagc	324
Qy	295	ggcatgggcggtggcggttctcggtccaacagcagcctggcggaacgcc	342
Db	325	cgattccgagatcctggacgaagatcatgttgtacgaacacgaacac	372

Search completed: June 28, 2002, 11:27:18
Job time: 3895 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 28, 2002, 10:21:28 ; Search time 1603.91 Seconds
(without alignments)
2877.941 Million cell updates/sec

Title: US-09-829-124-1
Perfect score: 342
Sequence: 1 atggactctatcggaacaa.....gcagcctggcggaacgcc 342

Scoring table: IDENTITY_NUC

Gap 10.0 ; Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues

Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_inv:*
15: em_gss_pln:*
16: em_gss_vit:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	43.8	12.8	526	12	AZ312601 1M0028104
2	41.8	12.2	537	10	BG930118 etshEST05
3	41.6	12.2	571	12	AQ919110 RPCI-23-2
4	41.2	12.0	505	9	BB867074
5	41.2	12.0	968	12	CNS03F00
6	40.8	11.9	425	10	B1895704
7	40.6	11.9	803	12	AZ101123
8	40.6	11.9	325	10	BG929756
9	40.6	11.9	559	12	AZ661219
10	40.4	11.8	585	10	B1832289
11	40.4	11.8	815	9	AU142497
12	40.4	11.8	850	10	BG417526
13	40	11.7	291	10	BF592443
14	39.8	11.6	605	12	AZ640388
15	39.2	11.5	507	10	BF199083
16	39.2	11.5	650	12	AZ337339
17	39	11.4	420	10	BG562218

C 18	39	11.4	821	10	BF570046
19	38.8	11.3	546	9	AJ273961
20	38.8	11.3	650	10	BM488443
C 21	38.6	11.3	440	10	BF414279
C 22	38.6	11.3	718	12	AZ972907
23	38.4	11.2	1006	12	CNS04SSP
24	38.2	11.2	541	12	B73805
25	38	11.1	330	10	BG560855
26	38	11.1	354	9	BB870603
C 27	38	11.1	501	9	AT366107
28	38	11.1	615	9	AW106297
29	38	11.1	617	9	BB619432
30	38	11.1	635	9	BB619096
31	38	11.1	818	10	BG297454
32	37.8	11.1	362	10	BF072844
33	37.8	11.1	389	12	AZ537043
34	37.8	11.1	689	10	BG855860
35	37.8	11.1	812	12	CNS03005
36	37.8	11.1	1042	12	CNS04NIV
37	37.6	11.0	840	9	AL669315
38	37.4	10.9	609	10	B1633649
39	37.4	10.9	822	10	BG299472
40	37.2	10.9	261	12	BH326016
C 41	37.2	10.9	348	9	AI440942
42	37.2	10.9	423	10	BM305606
43	37.2	10.9	440	10	B1895386
C 44	37.2	10.9	472	9	AW756485
C 45	37.2	10.9	473	9	AW277347

ALIGNMENTS

RESULT 1
AZ312601/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AZ312601 526 bp DNA linear GSS 29-SEP-2000
1M0028104R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
clone UUGC1M0028104 R, DNA sequence.
GSS.
house mouse.
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 526)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112, USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0028 row: 1 column: 04
Seq primer: CACACAGGAACACGCTATGACC
Class: plasmid ends
High quality sequence stop: 526.

FEATURES
Location/Qualifiers
1..526
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0028104"
/clone_lib="Mouse 10kb plasmid UUGC1M library"

/sex="Male"
 /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
 /note="Vector: PWD42nv; Purified genomic DNA from M.
 musculus C57BL/6J (male) was obtained from the Jackson
 Laboratory Mouse DNA Resource
 (http://www.jax.org/resources/documents/dnares/). The DNA
 was hydrodynamically sheared by repeated passage through a
 0.005 inch orifice at constant velocity. The sheared DNA
 was blunt end-repaired with T4 DNA polymerase and T4
 polynucleotide kinase. Adaptor oligonucleotides were
 ligated to the blunt ends in high molar excess. The
 adaptor DNA was purified and size-selected for a 9.5 to
 10.5 kb range using preparative agarose gel
 electrophoresis. Vector DNA was prepared from a derivative
 of pWD42 (g14732114|gb|AF129072.1), a copy-number
 inducible derivative of plasmid R1. The vector was ligated
 with adaptors complementary to the insert adaptors and
 purified. The sheared, adaptor mouse DNA was annealed to
 adaptor vector DNA, and transformed into
 chemically-competent E. coli XL10-Gold (Stratagene) cells
 and selected for ampicillin resistance."

[illegible]

RESULT. 2	EG930118	537 bp	mrna	linear	EST 31-DEC-2001
LOCUS	EG930118	etsHST0539	Eth11	Eimeria tenella	cDNA clone. etshk077 5', mrna
DEFINITION					sequence.
ACCESSION	EG930118				
VERSION	EG930118.1	GI:18003508.			
KEYWORDS					EST.
SOURCE					Eimeria tenella.
ORGANISM					Eimeria tenella
					Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.
REFERENCE					1 (bases 1 to 537)
AUTHORS					Ng, S.T., Jangli, M.S., Shirley, M.W., Tomley, F.M. and Wan, K.L.
TITLE					Comparative EST analyses provide insights into gene expression in two asexual developmental stages of Eimeria tenella
JOURNAL					Unpublished (2001)
COMMENT					Contact: Wan KL
					Centre for Gene Analysis and Technology
					Universiti Kebangsaan Malaysia
					43600 UKM Bangi, Selangor DE, Malaysia
					Tel: 6 03 8292997
					Fax: 6 03 8293249

```

Email: kluan@pkprisc.cc.ukm.my
PCR Primers
FORWARD: T3
BACKWARD: T7
Seq primer: SK.

FEATURES
    source
        1..537
            /organism="Eimeria tenella"
            /strain="Houghton"
            /db_xref="taxon:5802"
            /clone="etsHK077"
            /clone_lib="Eth11"
            /dev_stage="Sporozoite"
            /lab_host="XLI-Blue MRP'"
            /note="Vector: Lambda ZAPII; Site_1: EcoRI; Site_2: XhoI; Sporozoites were excysted in vitro from E. tenella H oocysts, purified by column chromatography and mRNA extracted using a FASTTRACK kit (Invitrogen). cDNA was synthesised and a Uni-ZAP XR library was constructed using cDNA synthesis kit, ZAP-cDNA synthesis kit and ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The library was amplified once through E. coli XLI-Blue MRP'."
        175 a 140 c 191 g 29 t 2 others

```

BASE COUNT	175 a	140 c	191 g	29 t	2 others
ORIGIN					
Query Match	12.2%;	Score 41.8;	DB 10;	Length 537;	
Best Local Similarity	48.2%;	Pred. No. 10;			
Matches 118;	Conservative 0;	Mismatches 127;	Indels 0;	Gaps 0;	
QY	37	aacctgcagacgatggcgctcggcctccacacacgagagactccagccagcagctcgccct	96		
Db	153	AGCTGAGAGAGAAAGCGGCAGCAGCAGCACTCTGGAGAGAGACGACGACGACGAG	212		
QY	97	tcggctggctccgagcagcagctggatcagttgctgcctatgttccatcatgatgatctg	156		
Db	213	GAGGCCGAGAGGAAGCAGCAGCAGCAAGCAGACTCGCGCGAGGAGTGTTCGGAAGCAG	272		
QY	157	caacagacgccaggcgagcgatgcaaatcagagagtgtgtgcaacacacacgcagaaacggt	216		
Db	273	CACAGAGAGTCCAGGGCCGCGCAGCAGCAGCAAGAGAGTCCAGAGAGAGAGACGACGAG	332		
QY	217	caacagagaagcctgagtcggttgacgcagatgctgtgcagatcgtgtgagctgatg	276		
Db	333	CAGAAGACTCTGACGACACGCGCAGCAGCAGCAAGACTCTGAGGAGAGACGACGACGAG	392		
QY	277	cgaaa	281		
Db	393	CAGCA	397		

RESULT	3	
AQ919110		
LOCUS		
DEFINITION		
ACCESSION		
VERSION		
KEYWORDS		
SOURCE		
ORGANISM		
REFERENCE		
AUTHORS		
TITLE		
JOURNAL		
COMMENT		

TITLE	JOURNAL	COMMENT
1. The Role of the Teacher in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 1, pp. 1-10
2. The Impact of Technology on the Classroom	Journal of Educational Research	1980, Vol. 83, No. 2, pp. 11-20
3. The Role of the Student in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 3, pp. 21-30
4. The Impact of the Teacher on the Student	Journal of Educational Research	1980, Vol. 83, No. 4, pp. 31-40
5. The Role of the Parent in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 5, pp. 41-50
6. The Impact of the Parent on the Student	Journal of Educational Research	1980, Vol. 83, No. 6, pp. 51-60
7. The Role of the Community in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 7, pp. 61-70
8. The Impact of the Community on the Student	Journal of Educational Research	1980, Vol. 83, No. 8, pp. 71-80
9. The Role of the Society in the Classroom	Journal of Educational Research	1980, Vol. 83, No. 9, pp. 81-90
10. The Impact of the Society on the Student	Journal of Educational Research	1980, Vol. 83, No. 10, pp. 91-100

The Institute of Physical and
1-7-22 Suehiro-cho, Tsurumi-ku

1. .571

/note="Organ: Kidney/Brain; Vector: pBACE3.6; Site_1: EcoRI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or brain genomic DNA was isolated and partially digested with a combination of EcoRI and EcoRI Methylase. Size selected DNA was cloned into the pBACE3.6 vector at the EcoRI sites. The ligation products were transformed into DH10B electrocompetent cells (BRL Life Technologies)."

222 a 112 c 142 g 95 t
UNIB electrocompetent cells (BRL Life Technologies).

12.2%; Score 41.6; DB 12; Length 571;
48.0%; pred No 12.

```

further details.
e mouse tissues.
Location/Qualifiers
1. 505
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RIKEN full-length enriched, pooled cell lines"
/note="pooled cell lines : (cell_line=CRL-1751 WEHI 164),
(cell_line=CRL-2116 JC), (cell_line=RCB-0035 WEHI-3),
(cell_line=RCB-0464 Meth-A), (cell_line=RCB-0545 OHTA),
(cell_line=RCB-0559 K-1 Fl), (cell_line=RCB-1283 B16
melanoma), (cell_type=B cells, cell_line=CRL-1702 WEHI 231
melanoma), (cell_type=Leydig cells, cell_line=CRL-2065 MLTC-1),
(cell_type=Nullipotent stem cell, cell_line=CRL-2070 NE),
(tissue_type=bladder, cell_line=RCB-0544 MBT-2),
(tissue_type=bone marrow, cell_type=stroma cell,
cell_line=CRL-2028 SR-4367), (tissue_type=colon,
cell_line=RCB-0549 CR-43), (tissue_type=kidney,

```

BASE COUNT	BC3H1)*	64 t
ORIGIN	121 a 161 c	159 g

7074.1 GI:17113284

	Matches	136;	Conservative	0;	Mismatches	158;	Indels	0;	Gaps	0;
QY	38	acctgcagcagatgggcctcagcaacagaggaactccagcagcagtcgcctt	97							
Db	51	AAC TG CAG CAG GAG C T C C T G G T C C T C A A G C A G C A G C A G C T C C A G A A G C A G C A G C T C C T G T	110							

Qy	98	cggctggtccagcagcagcgatgatgcatttgcctcatgatgaagtgc	157
Dd	111	TGCGGAGTTCACGAAGCAGCACCACTTTGCAGCGGGCAGCAGGATCCAGGTGCAGA	170
Qy	158	aacagaccaggcgagcgatcaaatcaggagtgaggacaacaaacccagaacggtc	217
Dd	171	AGCACCCTAACAGCAGCAGCAGAGATGCTGGCGGCTAAGAGGCCAGGCTGGAGCAGCAGG	230
Qy	218	aacaggaggcgctgagtcgttagcgcagatgctgatgcagatcgtagcgactgatgc	277
Dd	231	GGCAGCGGAGCAGCAGCGGCAGGAGGTGGAGAAACACGCGCTGGAGCAGCAGCTGC	290
Qy	278	agaaccaggcgcgccgcggcatggcggtgcccgttcggtccaacagcagcctgg	331
Dd	291	TCATCTCTGCGCAACAAGAGAGAAGCAAGAGATGCCATGCCAGCACCCAGG	344

RESULT 5
 CNS03FQQ/c

LOCUS
 DEFINITION CNS03FQQ 968 bp DNA linear GSS 17-MAY-2000
 Tetraodon nigroviridis genome survey sequence T7 end of clone
 021M15 of library G from Tetraodon nigroviridis , genomic survey
 sequence.

ACCESSION
 VERSION AL241091.1 GI:7961860
 KEYWORDS GSS; genome survey sequence.
 SOURCE Tetraodon nigroviridis.
 ORGANISM Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphea; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontidae; Tetraodonto.
REFERENCE
AUTHORS 1 (bases 1 to 968)
 Roest-Crollius,H., Jaillon,O., Dasilva,C., Fizames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Beirnot,A. and
 Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
JOURNAL
REFERENCE 2 (bases 1 to 968)
AUTHORS Roest-Crollius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
 Tetraodon nigroviridis DNA sequence
JOURNAL
REFERENCE 3 (bases 1 to 968)
AUTHORS Genoscope.
TITLE Direct Submission
SUBMITTED (12-APR-2000) to the EMBL/GenBank/DBJ databases
This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
<http://www.genoscope.cns.fr/Tetraodon>.
FEATURES
 Location/Qualifiers
 1..968
 /organism="Tetraodon nigroviridis"
 /db_xref="taxon:99883"
 /clone_021m15"
 /clone_lib="G"
 /note="Genoscope sequence ID : CDBG02IAG08LP1-end : T7"
BASE COUNT 214 a 188 c 292 g 218 t 56 others
ORIGIN

Qy	117	gctggatcagtgtgcctcatgttcatgatgtcgcaacagagccagggcagcga	176
Dd	289	GCACGACCAGGASCCMGACCMBSHYCCMGACCAGGAGGCCAASCCAGACCAGGA	230
Qy	177	tgc aaatcag agtg tgc caa cga aac acc ccaga acg gtc caa cag ga	224
Dd	229	SCCAGCACGAGCGACGACCAAGAACCCAGMCCACAASCAGACCAGGA	182
RESULT	6		
BI895704			
LOCUS	425 bp	mRNA	linear EST 15-OCT-2001
DEFINITION	EESTed23h03.y1 Eimeria tenella S5-2 cDNA Neg Selected Eimeria tenella cDNA 5', mRNA sequence.		
ACCESSION	BI895704		
VERSION	BI895704.1	GI:16138840	
KEYWORDS	EST.		
SOURCE	Eimeria tenella.		
ORGANISM	Eimeria tenella Eukaryota; Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae; Eimeria.		
REFERENCE	1 (bases 1 to 425)		
AUTHORS	Liberator,P., Diaz,C., Tang,K., Marra,M., Hillier,L., Kucaba,T., Martin,J., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Florence,N., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterson,R., Wilson,R. and Sibley,D. WashU-Merck Eimeria tenella project		
TITLE	Unpublished (1999)		
JOURNAL	Contact: David Sibley, Ph.D. WashU-Merck Eimeria tenella project		
COMMENT	Washington University School of Medicine 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA Tel: 314 286 1800 Fax: 314 286 1810 Email: est@watson.wustl.edu Contact David Sibley (toxest@orcim.wustl.edu) for further information relating to organism, libraries, or clone availability. Seq primer: -4ORP from Glbco High quality sequence stop: 391. Location/Qualifiers 1..425 /organism="Eimeria tenella" /strain="LS18" /db_xref="taxon:5802" /clone.lib="Eimeria tenella S5-2 cDNA Neg Selected" /dev_stage="Sporozoite stage" /lab_host="SOLR" /note="Vector: Bluescript SK+; Site_1: EcoRI; Site_2: XhoI ; Sporozoites were obtained from in vitro sporulated and excysted oocysts of E.tenella grown in chickens. CDNA was synthesized from poly mRNA using an oligo-dT primer containing a xhoI site. Following second strand synthesis, EcoRI adapters were ligated to the cDNA and products were size-selected on Sephadryl S500. The cDNA were ligated to EcoRI/XhoI prepared lambda ZapII (Stratagene). Clones were converted to phagemids by mass excision using ExAssist helper phage and E.coli SOLR cell (Stratagene). Clones were selected by negative hybridization against a pool of overrepresented ESTg (>-10, from 1682 previous reads). Insert sizes range from 1.2-2.9Kb. The library may contain a small percentage of host or bacterial contaminants."		
BASE COUNT	127 a	164 c	107 g 26 t
ORIGIN			1 others

	Query Match	12.0%	Score 41.2;	DB 12;	Length 968;
	Best Local Similarity 41.1%;	Pred. No. 17;			
	Matches 69;	Conservative 27;	Mismatches 72;	Indels 0;	Gaps 0;
QY	57	cgggcttcagcaacagaggaattccagccagcagtcgcttcgctggctccgagcagca	116		
		: : : : : : :			
Db	349	CGGGGWCAGGCGCCGACCCWMCSCMCCMGAAACCCAGMCCMSASCCMCCAGAA	290		

[illegible]

QY	10	atcggaacaacttttcgaatatatcgcaacctgcagacgatggcgatcgggcttcagcaa	69
DB	11		70
DB	12	AGCAGCAGCAGCAGTAGTAGCAGACGACGACGACGACGAGTAGCAGAAAGCAGAGAAG	71
QY	70	cacgaggactccagccagcagtcgcgcttcggctcggctccgagcagcagctggatcagttg	129
DB	71		130
DB	72	CAGAAGCAGAAGCAGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	131
QY	130	ctcgccatgttcatatgatgtctgtcaacagagccaggcgagcgatgcaaatcaggag	189
DB	131		190
DB	132	CAGCAGCAGAAGCAGAAGCAGAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	191
QY	190	tgtggcaacgaacaaacgcgacagtcgtcaacaggaaaggctcgttcggttgacgcagatg	249
DB	191		250
DB	192	CAGAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG	251
QY	250	ctgatgcagatcgtgatgcagctgtatgcagaacacag	285
DB	251		286
DB	252	CAGAAGCAGAAGCAGAAGCAGAAGCAGAAGCAGAAGCAGCAGCAGCAG	287
DB	253		288

[illegible]

REFERENCE	1 (bases 1 to 325)
AUTHORS	Ng, S.T., Jang, M.S., Shirley, M.W., Tomley, F.M. and Wan, K. L.
TITLE	Comparative EST analyses provide insights into gene expression two asexual developmental stages of <i>Eimeria tenella</i>
JOURNAL	Unpublished (2001)
COMMENT	Contact: Wan KL Centre for Gene Analysis and Technology Universiti Kebangsaan Malaysia 43600 UKM Bangi, Selangor DE, Malaysia Tel: 6 03 8292997 Fax: 6 03 8293249 Email: kiwan@prisc.cc.ukm.my PCR Primers FORWARD: T3 BACKWARD: T7

```

FEATURES             Seq primer: SK.
                     Location/Qualifiers
     1..325
         /organism="Eimeria tenella"
         /strain="Houghton"
         /db_xref="taxon:5802"
         /clone="etsHd066"
         /clone_lib="Eth11"
         /dev_stage="Sporozoite"
         /lab_host="XL1-Blue MRF/"
         /note="vector: Lambda ZAPII; Site_1: EcoRI; Site_2: XbaI"
     Sporozoites were excysted in vitro from E. tenella H
     oocysts, purified by column chromatography and mRNA
     extracted using a FASTTRACK kit (Invitrogen). cDNA was
     synthesised and a Uni-ZAP XR library was constructed
     using cDNA synthesis kit, ZAP-cDNA synthesis kit and
     ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The
     library was stored at -80°C.

```

BASE COUNT	72 a	97 c	102 g	54 t
ORIGIN				
Query Match	11.3%; Score 40.6; DB 10; Length 325;			
Best Local Similarity	51.5%; Pred. No. 17;			

of pWD42 (gii4732114|gb|AF129072.1), a copy-number inducible derivative of plasmid pL. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adapted mouse DNA was annealed to adapted vector DNA, and transformed into chemically-competent *E. coli* XL10-Gold (Stratagene) cells and selected for ampicillin resistance.

17 a	250 c	143 g	149 t
------	-------	-------	-------

BASE COUNT
ORIGIN

Query Match 11.9%; Score 40.6; DB 12; Length 559;
Best Local Similarity 51.1%; Pred. No. 20;
Matches 120; Conservative 0; Mismatches 114; Indels 1; Gaps 1;

[illegible]

Db	243	CAGCATCGGCAGCAGCATCGCAGCATTTGGCAGCGCAGGGGCGACGACATCGGC	189
RESULT	10		
BI832289			
LOCUS	BI832289	585 bp	linear
DEFINITION	603080137F1 NIH_MGC_119 Homo sapiens cDNA clone IMAGE:5171817 5', mRNA sequence.		
ACCESSION	BI832289		
VERSION	BI832289.1		
KEYWORDS	EST.		
SOURCE	human.		

mRNA sequence.
 accession B1832289
 version B1832289.1 GI:15943839
 keywords EST.
 source human.
 organism Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 585)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 contact: Robert Strausberg, Ph.D.
 comment

```

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LML)
DNA sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LML at:
http://image.llnl.gov
Plate: L1AM11427 row: j column: 10
High quality sequence start: 8
High quality sequence stop: 375.
Location/Qualifiers
1..585
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:5171817"
/clone_lib="NIH_MGC_119"
/tissue_type="medulla"
/lab_host="DH108"
/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;
Site_3: EcoRV (destroyed); RNA source normal medulla from
anonymous male age 27. Library is oligo-dT primed and
directionally cloned (EcoRV site is destroyed upon
cloning). Average insert size 1.3 kb, insert size range

```


	Matches	92;	Conservative	0;	Mismatches	86;	Indels	0;	Gaps	0;
QY	116	aqctggatcaq	tctctcgccat	tctcatgatgat	gctgcaacagaccagac	g	175			

Db 371 AGCACAATGACTTTCTGAAGCTGAGAACCATGCTCATGAAAGTGGAGAAATGAGGACATGA 430

Db 431 ATAAAGACCGAGATCTTGCTGGAAAAGAAGCTGAGCTCCGCCGATCGCAAGAGATGATTG 490

Db
491 CAAGGATCGAGCGGAGATGCAGATGCAGGCGGGGATGGCGATGCGGGGC 548

LOCUS	850 bp	mRNA	linear	EST 23-OCT-2003
BG417526				
LOCUS				
BG417526				
RESULT 12				

Accession	Sequence
BG417526	HVCDNAU013 (normal) Hordeum vulgare cDNA clone HVSMEK0017024f, mRNA

KEYWORDS
EST.
SOURCE
barley.
ORGANISM
Hordeum vulgare

REFERENCE 1 (bases 1 to 850)
; Triticeae; Hordeum.
; Triticaceae; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae
; Spemmatophytas; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae

WHEELER, E.; ARKADIEV, E.; CHIRIA, R.; CHOI, D. W.; FEILICH, R. D.; KLANIAN, P.; OTTO, C.; SIMONS, K.; ZHANG, D.; BEGUM, D.; FRISCH, D.; YU, Y.; HENRY, D.; PALMER, M.; RAMBO, T.; SIMMONS, J.; OATES, R. and MAIN, D.

JOURNAL COMMENT

100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4203

Total hq bases = 528
Seq primer: AATTAACCCCTCACTAAAGG
High quality sequence stop: 933

```
source
1. .850
/organism="Hordeum vulgare"
/cultivar="Morex"
```

```

/clone="HVSMEK0017024f"
/clone_lib="Hordeum vulgare testa/pericarp EST library
HVCNA0013 (normal)"

```

```

/lab_host="TJCI21"
/note="Vector: lambdaZAP; Site_1: EcoR1; Site_2: Xho1;
Plants were raised from seeds in a Controlled Environments

```

testa and pericarp were dissected from developing kernels at Washington State University, Pullman, WA (Kannangara, von Wettstein). Total RNA was prepared, poly(A⁺ RNA was

phagemids in the TJ Close lab at the University of California, Riverside (Akhunov, Chin, Choi, Close, Fenton, *et al.* 2000). The phagemids were in vivo excised to give pBluescript SK(-) cDNA

picked at the Clemson University Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins and Wing). Plasmid DNA preparations, DNA sequencing and sequence analysis were

, Rambo, Mami). The sequence has been trimmed to remove

vector sequence and contains a minimum of 100 bases of phred value 20 or above. For more details on library preparation and sequence analysis see <http://www.genome.clemson.edu/projects/barley>. To order this clone see <http://www.genome.clemson.edu/orders> Also see Close TJ, Wing R, Kleinhofs A, Wise R (2001) Genetically and physically anchored EST resources for barley genomics. Barley Genetics Newsletter 31:29-30. (<http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html>)

```

BASE COUNT      249 a 184 c 230 g 187 t
ORIGIN

Query Match      11.8%; Score 40.4; DB 10; Length 850;
Best Local Similarity 51.1%; Pred. No. 25;
Matches 95; Conservative 0; Mismatches 91; Indels 0; Gaps 0;

QY 152 tgcgaacagagccgagcgatgcaaatcaggagtggtgcaacgaacccgcaga 211
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 177 TGAAGCCACAGAACGAGCGCATTCGGCATTCGGAGTGGCGGAACGGCGACACTCACA 236

QY 212 acggtcaacaggaagcctgagtccttgacgcagatgctgagatgctgacgc 271
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 237 GCCTGTGCAGGTGGACCCCGCTTCACGTACAGGACCTCCAGGTCTATGACGAACA 296

QY 272 tgatgcagaaccagggcgccgcatggcggtggtggttcggtcaacagcgcctgg 331
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 297 ACTTCAGACGGTCTCGCGCGGGAGGATTCCGGTCCGGTACGACGGCTTCTTGGCGG 356

QY 332 gggcca 337
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 ATGGCA 362

```

```

RESULT 13
LOCUS      BF592443
DEFINITION 7156C04.x1 NCI_CGAP Br16 Homo sapiens cDNA clone IMAGE:338694 3'
            similar to contains element MSRI MSRI repetitive element ;, mRNA
            sequence.
ACCESSION  BF592443
VERSION     BF592443.1 GI:11684767
KEYWORDS   EST.
SOURCE     human.
ORGANISM   Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE  1 (bases 1 to 291)
            NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
            National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
            Unpublished (1997)
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-remail.nih.gov
            Tissue Procurement: Dennis Sgroi, M.D., Kristina Cole, M.D., Ph.D.
            student, Michael R. Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: David B. Krizman, Ph.D.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone Distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL, send email to:
            info@image.llnl.gov
            Seq primer: -40RP from Gibco.
            Location/Qualifiers
                1..291
                /organism="Homo sapiens"
                /db_xref="taxon:9606"
                /clone="IMAGE:338694"
                /clone_lib="NCI_CGAP_Br16"
                /sex="female"
                /tissue_type="lobular carcinoma in situ"
                /dev_stage="adult"
                /lab_host="DH10B"

```

/note="Organ: breast; Vector: pAMP1; mRNA made from breast carcinoma tissue, cDNA made by oligo-dT priming. Directionally cloned. Size-selected on agarose gel, average insert size 400 bp. Primary library, non-amplified."

```

BASE COUNT      86 a 43 c 139 g 23 t
ORIGIN

Query Match      11.7%; Score 40; DB 10; Length 291;
Best Local Similarity 51.7%; Pred. No. 23;
Matches 91; Conservative 0; Mismatches 85; Indels 0; Gaps 0;

```

```

QY 111 gcaacagctgacgtgctgctgccatgttcacatgatgatgctgcaacagagccagg 170
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 69 GAAGCAGGAGGAGGATCGGGAAGCAGGAGGAGGATGGGGAGGAGGAGGAGGAGAT 128

QY 171 -cagcgtgcaaatcaggagtggtgcaacgaacacccgacacggtcaacaggaagcct 230
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 129 GGGGAGCAGGAGGAGGATCGGGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGACCA 188

QY 231 gactccgttgacgcagatgctgacgatcgtgacgtgacgtgacgtgacgaaccagg 286
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 189 GATGCTGAACGAGGAGGAGGATCGGGAAGCAGGAGGAGGAGGAGGAGGAGGAGGAGG 244

```

```

RESULT 14
LOCUS      AZ640388
DEFINITION 1M0502E21F Mouse 10kb plasmid UUGCLM library Mus musculus genomic
            clone UUGCLM0502E21 F, DNA sequence.
ACCESSION  AZ640388
VERSION     AZ640388.1 GI:11763788
KEYWORDS   GSS.
SOURCE     house mouse.
ORGANISM   Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE  1 (bases 1 to 605)
            Dunn, D., Aoyagi, A., Barber, M., Beacorn, T., Duval, B., Hamil, C.,
            Islam, H., Longacre, S., Mahmoud, M., Meenen, E., Pedersen, T., Rellly
            M., Rose, M., Rose, R., Stokes, R., Tingey, A., von Niederhausen, A.
            and Wright, D., Weiss, R.
            Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
            Unpublished (2000)
            Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000 Std Error: 0.00
            Plate: 0502 row: E column: 21
            Seq primer: CGTTGTAACGACGCGCCAGT
            Class: plasmid ends
            High quality sequence stop: 605.
            Location/Qualifiers
                1..605
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone="UUGCLM0502E21"
                /clone_lib="Mouse 10kb plasmid UUGCLM library"
                /sex="Male"
                /lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
                /note="Vector: pWD42hv; Purified genomic DNA from M.
                musculus C57BL/6J (male) was obtained from the Jackson
                Laboratory Mouse DNA Resource
                (http://www.jax.org/resources/documents/dnares/). The DNA
                was hydrodynamically sheared by repeated passage through a

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2002, 10:16:03 ; Search time 23.71 Seconds
(without alignments)
534.054 Million cell updates/sec

Title: US-09-829-124-2
Perfect score: 584
Sequence: 1 MDSIGNFNSNIGNLQTMGIG.....OGGAGMGGGSVNSSLGNA 114

Scoring table: BLOSUM62
Gapop 10.0 ; Gapext 0.5

Searched: 747574 seqs, 111073796 residues
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_032802:*

1:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22:	/SIDS1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	584	100.0	114	22	AAE10807
2	111	19.0	344	18	AAW06600
3	111	19.0	344	19	AAW75865
4	111	19.0	344	19	AAW62457
5	111	19.0	344	19	AAW61116
6	111	19.0	344	20	AAW87641
7	111	19.0	344	21	AAW71099
8	111	19.0	344	21	AAW84860
9	111	19.0	344	22	AAE06712
10	111	19.0	385	15	AAW45751
11	111	19.0	385	18	AAW06598

12	111	19.0	403	19	AAW75863
13	111	19.0	403	19	AAW62455
14	111	19.0	403	19	AAW61114
15	111	19.0	403	20	AAW87639
16	111	19.0	403	21	AAW71093
17	111	19.0	403	21	AAW84854
18	111	19.0	403	22	AAE06710
19	106	18.2	318	22	ABW68750
20	96.5	16.5	1366	22	ABW68921
21	96.5	16.5	1594	22	ABW63688
22	94	16.1	20	19	AAW75867
23	94	16.1	20	19	AAW62459
24	94	16.1	20	19	AAW61118
25	94	16.1	20	20	AAW87643
26	94	16.1	20	21	AAW71101
27	94	16.1	20	21	AAW84862
28	91.5	15.7	424	20	AAW97851
29	91.5	15.7	424	21	AAW71098
30	91.5	15.7	424	21	AAW84859
31	90	15.4	2328	22	ABW71657
32	90	15.4	3190	22	ABW84634
33	90	15.4	3275	22	ABW70437
34	88	15.1	800	16	AAW84883
35	88	15.1	2414	16	AAW84882
36	88	15.1	2414	19	AAW40057
37	87	14.9	724	22	ABW65778
38	87	14.9	1778	22	ABW60081
39	85	14.6	147	22	ABW69847
40	84.5	14.5	797	22	ABW62042
41	84.5	14.5	915	22	ABW65275
42	84	14.4	447	20	AAW96260
43	84	14.4	447	21	AAW71094
44	84	14.4	447	21	AAW84855
45	84	14.4	926	22	ABW65135

ALIGNMENTS

RESULT 1

AAE10807
ID AAE10807 standard; Protein; 114 AA.
XX
AC AAE10807;
XX
DT 18-DEC-2001 (first entry)
XX
DE Xanthomonas campestris hypersensitive response elicitor protein.
XX
KW Plant pathogen; hypersensitive response elicitor; HRE; stress resistance;
KW disease resistance; Erwinia amylovora hypersensitive response elicitor;
KW harpin; signal transduction; growth enhancement; insect control;
KW virucide; fungicide; antibacterial.
XX
OS Xanthomonas campestris.
XX
PN WO200170988-A2.
XX
PD 27-SEP-2001.
XX
PF 19-MAR-2001; 2001WO-US08728.
XX
PR 23-MAR-2000; 2000US-191649P.
XX
PR 01-DEC-2000; 2000US-250710P.
XX
PA (EDEN-) EDEN BIOSCIENCE CORP.
XX
PI Song X, Fan H, Wei Z;
XX
DR WPI; 2001-590177/66.
XX
DR N-PSDB; AAD18127.
XX
PT New plant pathogen hypersensitive response elicitor-receptor protein

Erwinia amylovora
Erwinia amylovora
Hypersensitive res
A hypersensitive r
Erwinia amylovora
A hypersensitive r
Erwinia amylovora
Drosophila melanog
Drosophila melanog
Xanthomonas campe
Xanthomonas campe
Hypersensitive res
A hypersensitive r
Xanthomonas campe
Hypersensitive res
Hypersensitive res
Pseudomonas syring
A hypersensitive r
Drosophila melanog
Amino acid sequenc
Drosophila melanog
Transcription fact
Transcription fact
Cellular transcrip
Drosophila melanog
Drosophila melanog
Drosophila melanog
Drosophila melanog
Hypersensitive res
Erwinia amylovora
A hypersensitive r
Drosophila melanog

PT isolated from plants, which upon silencing is used to study plant
PT signal transduction pathways leading to disease resistance and growth
XX enhancement

PS Disclosure: Page 18-19; 78pp; English.

XX The present invention relates to an isolated protein which serves as a
CC receptor in plants for plant pathogen hypersensitive response elicitors
CC (HRE). The plants are made resistant to infection by viruses, bacteria
CC and fungi and are imparted with resistance against environmental stress
CC and insects through HRE treatment. The protein is useful for
CC understanding the harpin (Brwinia amylovora hypersensitive response
CC elicitor) induced signal transduction pathway in plants. The protein is
CC useful for studying the downstream components of signal transduction
CC pathway in plants which eventually leads to disease resistance, growth
CC enhancement, insect control and stress resistance. Imparting disease
CC resistance to plants through HRE treatment has the potential to treat
CC previously untreatable diseases, treating diseases systemically which
CC might not be treated separately due to cost and avoids the use of
CC infectious agents or environmentally harmful materials. By HRE treatment
CC enhanced plant growth is achieved which includes greater yield,
CC increased quantity of seeds produced, increased percentage of seeds
CC germinated, increased plant size, greater biomass, more and bigger
CC fruits, etc. which results in economic benefit to cultivators. Greater
CC yield, increased size and enhanced biomass allow greater revenue
CC generation from the given plot of plant. The present sequence is
CC Xanthomonas campestris pv. pelargonii hypersensitive response elicitor
CC protein.

SQ Sequence 114 AA;

Query Match 100.0%; Score 584; DB 22; Length 114;
Best Local Similarity 100.0%; Pred. No. 4.8e-57;
Matches 114; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MDSIGNFNSNIGNTMTGIGPOOHEDSSQSPSAGSEQDLQLAMFIMMQLQSGSDA 60
|||||

.Db 1 mdsignfnsnigntmtgigpqhdsqqspgsaqeqldqlamfimmqlqsggsda 60
|||||

QY 61 NQPCNEPONGQQGLSPLTOMLMOIVMQLMONGGAGMGGSVNSSLGNA 114
|||||

.Db 61 nqcgneqpnqgqeglspltgmlmqimqlmqngggagmgggsvnsslgnna 114
|||||

RESULT * 2

AAW06600
ID AAW06600 standard; Protein; 344 AA.

XX AAW06600;

XX 30-MAR-1997 (first entry)

XX Hypersensitive response elicitor protein.

XX Hypersensitive response; elicitor; Pseudomonas solanacearum; plant;
KW disease-resistance; Escherichia coli; infiltration; virus;
KW bacterium; fungus; pathogen; biological control agent.

XX Pseudomonas solanacearum.

XX WO9639802-A1.

XX 19-DEC-1996.

XX 05-JUN-1996; 96WO-US08819.

XX 07-JUN-1995; 95US-0475775.

XX (CORR) CORNELL RES FOUND INC.

XX Beer SV, Wei 2;

XX

DR WPI: 1997-051614/05.
XX N-PSDB; AAT49316.

PT Imparting pathogen resistance to plants - with hypersensitive
PT response elicitor polypeptide or protein

PS Claim 13; Page 50-51; 69pp; English.

XX This sequence represents a hypersensitive response elicitor from
CC Pseudomonas solanacearum. The elicitor may be used in a new method
CC for imparting pathogen resistance to plants, by application of the
CC elicitor in a non-infectious form to plant cells, by spraying,
CC injection, leaf abrasion, or plant infection with recombinant
CC bacteria (non-infectious to the host plant, e.g. Escherichia coli)
CC expressing the elicitor as a biological control agent, to allow
CC recombinant protein infiltration into the plant. The method
CC confers virus, bacterium or fungus disease-resistance on crops and
CC ornamental plants.

SQ Sequence 344 AA;

Query Match 19.0%; Score 111; DB 18; Length 344;
Best Local Similarity 25.8%; Pred. No. 0.00045;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

QY 3 SIGNFSNIGNTMTGIGPOOHEDSSQSPSAGSE-----QQLDQ 42
|||

.Db 58 sagngtntgnapakdgnanagandpskndpsksgqpsanktgnvddannqdpmqalmq 117
|||

QY 43 LLAMFIMMM-----LQSQSGSDANQECG--NEQPQNGQOEGLSPLTOMLMOIVMQLMONG- 94
|||

.Db 118 lledivkilklaahmqpggndkngvggagangakagagggglaealqeieqlaglgggg 177
|||

QY 95 -----QGGAGMGGSVNSSLGNN 113
|||

.Db 178 agagagagvggagagadggsgagaggangadggn 212
|||

RESULT 3

AAW75865

ID AAW75865 standard; Protein; 344 AA.

XX AAW75865;

XX 07-DEC-1998 (first entry)

XX Pseudomonas solanacearum hypersensitive response elicitor (HRE).

XX Hypersensitive response elicitor; HRE; insect resistance;
KW biological control; transgenic plant.

XX Pseudomonas solanacearum.

XX WO9837752-A1.

XX 03-SEP-1998.

XX 26-FEB-1998; 98WO-US03604.

XX 28-FEB-1997; 97US-0039226.

XX (CORR) CORNELL RES FOUND INC.

XX Wei 2, Zitter TA;

XX WPI: 1998-495374/42.

XX N-PSDB; AAV54609.

XX Use of hypersensitive response elicitor polypeptide - for
PT application to plants or seeds or transgenic plants or seeds for the
PT control of insects.

C This is the deduced amino acid sequence of the hypersensitive

Best Local Similarity 25.8%; Pred. No. 0.00045;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

QY 3 SIGNFNSNIGNLQTM-----GIGPQOHHEDSSQQSPSAGSE-----QQLDQ 42
DB 58 sagngtntgnapakdgnanagandpskndpskspqqsantgnvddannqpmqalmq 117
QY 43 LLAMFIMM-----LQOSGSDANQECG--NEQPONGQOEGLSPLTQMLQIVMQLMQN- 94
DB 118 lledlvkllaalhmqqpgndkngvggagangakgagggglaealqeieqilaqlgggg 177
QY 95 -----QGGAGMGGGSVNSLGGN 113
DB 178 agagaggvggagadggsgaggaggangadggn 212

RESULT 8

AAE06712
ID AAE06712 standard; Protein; 344 AA.

AC AAE06712;

XX 08-AUG-2000 (first entry)

XX A hypersensitive response elicitor protein.

XX Hypersensitive response; insect control; disease resistance;

KW Hypersensitive response elicitor; plant growth; vegetable; crop;

XX ornamental plant.

XX Pseudomonas solanacearum.

XX WO200020452-A2.

XX 13-APR-2000.

XX 05-OCT-1999; 99WO-US23181.

XX 05-OCT-1998; 98US-0103050.

XX (EDEN-) EDEN BIOSCIENCE CORP.

XX Wei Z, Fan H, Niggemeyer JL;

XX WPI; 2000-303745/26.

XX N-PSDB; AAA14944.

XX Hypersensitive response elicitor polypeptides useful for imparting
enhanced growth, disease resistance and insect resistance to plants,
especially vegetables and ornamental flowers.

XX Disclosure; Page 28-29; 100pp; English.

XX The present sequence represents a hypersensitive response elicitor
polypeptide. The specification describes hypersensitive response
elicitor polypeptide fragments, which do not elicit a hypersensitive
response. Instead, the proteins impart disease resistance to plants,
enhance plant growth, and/or control insects. The polypeptide
fragments may be used to these properties to plants. The plants which
may be treated in this way include vegetables, crops and ornamental
plants such as alfalfa, rice, wheat, barley, rye, cotton, sunflower,
peanut, corn, potato, sweet potato, bean, pea, chickory, lettuce,
endive, cabbage, brussel sprout, beet, parsnip, turnip, cauliflower,
broccoli, radish, spinach, onion, garlic, eggplant, pepper, celery,
carrot, squash, pumpkin, zucchini, cucumber, apple, pear, melon,
citrus, strawberry, grape, raspberry, pineapple, soybean, tobacco,
tomato, sorghum or sugarcane, Arabidopsis thaliana, Saintpaulia,
petunia, pelargonium, poinsettia, chrysanthemum, carnation or zinnia.

XX Sequence 344 AA;

Query Match

19.0%; Score 111; DB 21; Length 344;

Best Local Similarity 25.8%; Pred. No. 0.00045;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

QY 3 SIGNFNSNIGNLQTM-----GIGPQOHHEDSSQQSPSAGSE-----QQLDQ 42
DB 58 sagngtntgnapakdgnanagandpskndpskspqqsantgnvddannqpmqalmq 117
QY 43 LLAMFIMM-----LQOSGSDANQECG--NEQPONGQOEGLSPLTQMLQIVMQLMQN- 94
DB 118 lledlvkllaalhmqqpgndkngvggagangakgagggglaealqeieqilaqlgggg 177
QY 95 -----QGGAGMGGGSVNSLGGN 113
DB 178 agagaggvggagadggsgaggaggangadggn 212

RESULT 9

AAE06712
ID AAE06712 standard; Protein; 344 AA.

AC AAE06712;

XX 16-OCT-2001 (first entry)

XX Pseudomonas solanacearum hypersensitive response elicitor protein.

KW Hypersensitive response elicitor; oomycete; transgenic plant; infection;
gene therapy; crop loss; antifungal.

XX Pseudomonas solanacearum.

XX WO200155347-A1.

XX 02-AUG-2001.

XX 26-JAN-2001; 2001WO-US02579.

XX 26-JAN-2000; 2000US-0178565.

XX (CORR) CORNELL RES FOUND INC.

XX Beer SV, Bauer DW;

XX WPI; 2001-488791/53.

XX N-PSDB; AAD12808.

XX New chimeric gene, useful for controlling plant-pathogenic fungi and
producing oomycete-resistant transgenic plants, comprises first DNA
encoding hypersensitive response elicitor, promoter and regulatory
region.

XX Claim 20; Page 16-17; 72pp; English.

XX The invention relates to a chimeric gene that includes a first DNA
molecule encoding a hypersensitive response elicitor protein or
polypeptide, promoter operably linked 5' to the first DNA molecule
to induce transcription of the first DNA molecule in response to
activation of the promoter by an oomycete and a 3' regulatory region
to a transgenic plant resistant to disease resulting from oomycete
infection, the transgenic plant including the chimeric gene, wherein
the promoter induces transcription of the first DNA molecule in
response to infection of the plant by an oomycete. The chimeric gene
is used in gene therapy. The chimeric gene is useful as an effective
and safe means of controlling plant-pathogenic fungi, particularly
oomycetes, which are responsible for major crop loss and is also useful
for producing transgenic plants of the invention. The present sequence
is Pseudomonas solanacearum hypersensitive response elicitor protein.

XX Sequence 344 AA;

Query Match

19.0%; Score 111; DB 22; Length 344;

Best Local Similarity 25.8%; Pred. No. 0.00045;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

QY 3 SIGNFNIGNLQTM-----GIGPQOHHEDSSQOQSPSAGSE-----QQLDQ 42
Db 58 sagntntgnapakdnaganandpskndpsksqapqsanktgnvddannqdpmqalmq 117
QY 43 LLAMFIMM-----LQSQSGSDANQEGC--NEQPQNGQOGLSPLTQMLMIVMQLMN- 94
Db 118 lledlvklkaalhmqqgndkngvggagkakgagggglaealqeieqlaqlggg 177
QY 95 -----QGGAGMGGGGVSNSLGN 113
Db 178 agagagggvggagdgagdgsgagagagadgn 212

RESULT 10
AAR45751
ID AAR45751 standard; Protein; 385 AA.
AC AAR45751;
XX
DT 11-JUL-1994 (first entry)
XX
DE Erwinia amylovora harpin.
XX
KW Harpin; hypersensitive response elicitor; HR-elicitor; fire blight;
KW Rosaceae; apple; pear; phytopathogenic bacteria; defence reaction;
KW hypersensitive reaction and pathogenicity; hrpN; gene cluster; ds.
XX
OS Erwinia amylovora.
XX
FH Key Location/Qualifiers
FT Misc-difference 372 /note= "corresponds to CAG codon"
FT Misc-difference 373 /note= "corresponds to CAC codon"
XX
PN WO9401546-A.
XX
PD 20-JAN-1994.
XX
PF 30-JUN-1993; 93WO-0506243.
XX
PR 01-JUL-1992; 92US-0907935.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
PI Bauer DW, Beer SV, Collmer A, He S, Laby R, Wei Z;
XX
DR WPI; 1994-035054/04.
DR N-PSDB; AAR55751.
XX
PT Hypersensitive response elicitor protein derived from Erwinia
PT amylovora - and DNA encoding it, useful for developing harpin
PT inhibitors to prevent e.g. fire blight of fruit
XX
PS Claim 6; Page 27-28; 47pp; English.
XX
CC The hrpN gene was isolated from E.amylovora using a 48-fold
CC degenerate oligonucleotide probe corresponding to amino acids 9-15
CC at the N-terminus of harpin. The 44kD protein encoded by the hrpN
CC gene is a hypersensitive response elicitor protein. The harpin is
CC thought to be an archetype for HR elicitors from phytopathogenic
CC bacteria.
XX
SQ Sequence 385 AA;

Query Match 19.0%; Score 111; DB 15; Length 385;
Best Local Similarity 31.5%; Pred. No. 0.00052;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

QY 5 GNNFSGNIGNL---OTMGI-GPQOHHEDSSQOQSPS-AGSEQQLDQLLAMFIMMLOQS---Q 56
Db 118 gnnstttnspldgaiginstsqnddstsgtdstsdspmqgllkmfseim--qslfgd 175
QY 57 GSDANQ--ECGNEQPQNGQOQ-----EGLSPLTQMLMIVMQLMNQGGAGMGGGGVNSL 110
Db 176 gqdggtqgsssggkqpteqeqnaykkgvtdalsgImngnqlsqllngpglggggggnagtgl 235
QY 111 GGNA 114
Db 236 dgss 239

RESULT 11
AAW06598
ID AAW06598 standard; Protein; 385 AA.
XX
AC AAW06598;
XX
DT 30-MAR-1997 (first entry)
XX
DE Hypersensitive response elicitor protein.
XX
KW Hypersensitive response; elicitor; Erwinia amylovora; plant;
KW disease-resistance; Escherichia coli; infiltration; virus;
KW bacterium; fungus; pathogen; biological control agent.
XX
OS Erwinia amylovora.
XX
PN WO9639802-A1.
XX
PD 19-DEC-1996.
XX
PF 05-JUN-1996; 96WO-US08819.
XX
PR 07-JUN-1995; 95US-0475775.
XX
PA (CORR) CORNELL RES FOUND INC.
XX
PI Beer SV, Wei Z;
XX
DR WPI; 1997-051614/05.
DR N-PSDB; AAT49314.
XX
PT Imparting pathogen resistance to plants - with hypersensitive
PT response elicitor polypeptide or protein
XX
PS Claim 7; Page 46-47; 69pp; English.
XX
CC This sequence represents a hypersensitive response elicitor from
CC Erwinia amylovora, with a mol.wt. of 37,000. The elicitor has a pI
CC of 4.3, thermostability at 100 deg C for at least 10 min, and
CC contains no cysteine. The elicitor may be used in a new method for
CC imparting pathogen resistance to plants, by application of the
CC elicitor in a non-infectious form to plant cells, by spraying,
CC injection, leaf abrasion, or plant infection with recombinant
CC bacteria (non-infectious to the host plant, e.g. Escherichia coli)
CC expressing the elicitor as a biological control agent, to allow
CC recombinant protein infiltration into the plant. The method
CC confers virus, bacterium or fungus disease-resistance on crops and
CC ornamental plants.
XX
SQ Sequence 385 AA;

Query Match 19.0%; Score 111; DB 18; Length 385;
Best Local Similarity 31.5%; Pred. No. 0.00052;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

QY 5 GNNFSGNIGNL---OTMGI-GPQOHHEDSSQOQSPS-AGSEQQLDQLLAMFIMMLOQS---Q 56
Db 118 gnnstttnspldgaiginstsqnddstsgtdstsdspmqgllkmfseim--qslfgd 175

Db 176 qdgdtgsssggkqpteqeqnkykavtdalslmgnglsqllnggllgggggnagtgl 235

QY 111 GGNA 114
|::

Db 236 dqss 239

RESULT 13
AAW62455
ID AAW62455 standard; Protein; 403 AA.

DT 09-NOV-1998 (first entry)

DE *Erwinia amylovora* hypersensitive response elicitor (HRE).

XX Hypersensitive response elicitor; HRE; growth; transgenic plant.

XX
OS Erwinia amylovora.

PN WO9832844-A1.

XX
PD 30-JUL-1998.

XX
PF
27-JAN-1998;XX
PR 27-JAN-1997;XX
PA (CORR) CORNXX
PI Beer SV, OiuXX
DR WPI; 1998-42

DR N-PSDB; AAV39
XX

Method for eliciting response elicited

Dbb 176 gdgdtqgsssgkaptgeeqnaykkvtdalsglmgnglsqllgn9glg9gg9nagtgl 238

Oy 111 CGNA 114
l::

Dbb 236 dgss 239

RESULT i4
AAW61114
ID AAW61114 standard; Protein; 403 AA.
XX AC
XX AAW61114;
DT DT 26-OCT-1998 (first entry)
XX
DE DE Hypersensitive response elicitor protein (39 kDa).
XX
KW KW Hypersensitive response elicitor; transgenic plant; seed;
KW KW pathogen resistance; disease resistance; crop protection.
XX
OS OS Erwinia amylovora.
XX
XX XX
FN FN W09824297-A1.
XX
PD PD 11-JUN-1998.
XX
PF PF 04-DEC-1997; 97WO-US22629.
XX
PR PR 05-DEC-1996; 96US-0033230.
XX
PA PA (CORR) CORNELL RES FOUND INC.
XX
PI PI Beer SV, Qiu D, Wei Z;
XX
DR DR WPI; 1998-332931/29.
DR N-PSDB; AAV36428.
XX
PT PT
PT Imparting pathogen resistance to plants - by applying a
PT hypersensitive response elicitor polypeptide to seeds
XX
PS Disclosure; Page 18-20; 85pp; English.
XX
CC This is the 39 kDa hypersensitive elicitor (HRE) protein of
CC Erwinia amylovora. It is heat stable at 100 degC for at least 10
CC min, has a pI of approximately 4.3, and contains substantially no
CC cysteine. The invention relates to methods of imparting
CC hypersensitive response induced resistance to plants by treatment
CC of seeds. Isolated HRE proteins can be applied to seeds as a means
CC of imparting pathogen resistance to plants grown from the seeds.
CC Alternatively, bacteria containing the gene encoding the HRE can be
CC applied to the plant seeds, or transgenic plant seeds containing a
CC DNA molecule encoding an HRE polypeptide or protein are used. HRE
CC polypeptide sequences from Erwinia chrysanthemi, Erwinia amylovora,
CC Pseudomonas syringae, Pseudomonas solanacearum, Xanthomonas
CC campestris pv. glycines and Xanthomonas campestris pelargonii (see
CC AAW61113-18) are provided. The methods can impart pathogen
CC resistance without using agents which are harmful to the
CC environment or pathogenic to the plant seed being treated, or to
CC adjacent plants. E. amylovora causes disease in apple or pear
CC but not tomato. However, it elicits a hypersensitive response in
CC tomato. Thus, E. amylovora can be applied to tomato seeds to
CC impart pathogen resistance without causing diseases in plants of
CC that species.

XX SQ Sequence 403 AA;

Query Match 19.0%; Score 111; DB 19; Length 403;
Best Local Similarity 31.5%; Pred. No. 0.00055;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps

OY 5 GNFSNIGNL--QTGMI-GPOQHEDSSQGSFSG-AGSEQLDQLLMTIMMLOQS---Q.56
||| :: | : || :||| : : : : ||| || :|||

Db 176 gqdgsgssgkqpteqeqnaykkgvtdalsgldnglsqlngglgsgsgnagtgl 235
Qy 111 GGNA 114
Db 236 dgss 239

Search completed: June 28, 2002, 10:21:25
Job time: 322 sec

Mon Jul 1 09:28:18 2002

us-09-829-124-2.ra1

Page 1

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using sw model

Run on: June 28, 2002, 10:16:33 ; Search time 13.3 Seconds
(without alignments)
209.362 Million cell updates/sec

Title: US-09-829-124-2
Sequence: 1 MDSIGNFNIGNLQPMGIC.....QGGAGMGGSVNSLAGNA 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 04
Maximum Match 1004
Listing first 45 summaries

Database : Issued Patents.AA.*
1: /cgn2_5/ptodata/2/laa/5A-COMB.pep.*
2: /cgn2_5/ptodata/2/laa/5B-COMB.pep.*
3: /cgn2_5/ptodata/2/laa/5A-COMB.pep.*
4: /cgn2_5/ptodata/2/laa/5B-COMB.pep.*
5: /cgn2_5/ptodata/2/laa/5A-COMB.pep.*
6: /cgn2_5/ptodata/2/laa/5B-COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	111	19.0	344	1	US-08-891-254-7
2	111	19.0	344	2	US-08-819-539-7
3	111	19.0	344	2	US-09-030-270A-7
4	111	19.0	344	4	US-08-984-207-7
5	111	19.0	344	4	US-09-013-587-7
6	111	19.0	344	5	PCT-US96-08819-7
7	111	19.0	385	1	US-08-891-254-3
8	111	19.0	385	2	US-08-819-539-3
9	111	19.0	385	5	PCT-US96-08819-3
10	111	19.0	403	2	US-09-030-270A-3
11	111	19.0	403	2	US-08-851-376A-3
12	111	19.0	403	4	US-08-984-207-3
13	111	19.0	403	4	US-09-013-587-3
14	111	19.0	20	2	US-08-030-270A-10
15	111	19.0	20	4	US-08-984-207-10
16	94	16.1	20	4	US-09-013-587-10
17	94	16.1	20	4	US-09-013-587-2
18	91.5	15.7	424	4	US-08-227-536-2
19	88	15.1	2414	1	PCT-US95-04682-2
20	88	15.1	2414	5	US-08-120-927-2
21	88	15.1	2414	4	US-08-851-376A-2
22	84	13.1	739	4	US-08-851-376A-2
23	74.5	12.8	653	4	US-09-061-764A-15
24	74.5	12.8	653	4	US-08-789-333F-41
25	74.5	12.8	653	4	US-08-789-333F-41
26	73	12.5	124	4	US-08-789-333F-41
27	73	12.5	173	4	US-08-789-333F-42

28	73	12.5	903	2	US-08-853-310-2
29	72	12.3	369	2	US-08-991-300-2
30	71.5	12.2	485	2	US-08-749-391-2
31	71.5	12.2	485	4	US-09-390-200-2
32	70.5	12.1	765	1	US-08-195-152-2
33	70	12.0	795	1	US-07-716-827C-5
34	67.0	12.0	2703	1	US-08-185-432-19
35	67	11.6	599	4	US-09-228-986-117
36	66.5	11.5	431	4	US-09-115-446-2
37	66.5	11.4	485	1	US-08-924-875-1
38	66.5	11.4	485	1	US-08-924-875-1
39	66.5	11.4	485	1	US-08-478-676-1
40	66	11.3	816	2	US-08-785-310A-8
41	66	11.3	816	2	US-08-816-693A-53
42	66	11.3	816	3	US-08-885-291-53
43	66	11.3	816	4	US-09-496-672-53
44	66	11.3	1004	4	US-08-916-352-2
45	65.5	11.2	255	4	US-09-342-084-11

ALIGNMENTS

RESULT 1
US-08-891-254-7
: Sequence 7, Application US/08891254
: Patent No. 5776889
: GENERAL INFORMATION:
: APPLICANT: Wei, Zhong-Min
: TITLE OF INVENTION: Hypersensitive Response
: NUMBER OF SEQUENCES: 9
: COUNTRY: U.S.A.
: ADDRESS: Naxon Hargrave, Devans & Doyle
: STREET: Naxon Square, P.O. Box 1051
: CITY: Rochester
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 14603
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: PUBLICATION NUMBER: US/08/891,254
: FILING DATE: 10-27-1997
: CLASSIFICATION: C12N 1/00
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/475,775
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldman, Michael L.
: REGISTRATION NUMBER: 30,727
: REFERENCE/DOCKET NUMBER: 14603/10050
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (716) 263-1304
: TELEFAX: (716) 263-1600
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 344
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
US-08-891-254-7

Query Match 19.0% Score 111; DB 1; Length 344;
Best Local Similarity 25.8%; Pred. No. 7.7e-05;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

```

Db 178 AGAGGAGGCGGGAGGAGGCGGGAGGAGGAGCGGN 212

RESULT 3
: Sequence 7, Application US/09030270A
: Patent No. 5977060
: GENERAL INFORMATION:
: APPLICANT: Zitter, Thomas A.
: APPLICANT: Wei, Zhong-Min
: TITLE OF INVENTION: INSECT CONTROL WITH A
: TITLE OF INVENTION: HYPERSENSITIVE RESPONSE ELICITOR
: NUMBER OF SEQUENCES: 0
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Nixon, Haigrave, Devans & Doyle LLP
: STREET: P.O. Box 1051, Clinton Square
: CITY: Rochester
: STATE: New York
: COUNTRY: U.S.A.
: ZIP: 14603
: COMPUTER READABLE FORM:
: DISC: 5.25 inch floppy disk
: MEDIUM TYPE: 5.25 inch floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/030.270A
: FILING DATE:
: CLASSIFICATION: 514
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/039,226
: FILING DATE: 28 FEB 1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Goldman, Michael L.
: REGISTRATION NUMBER: 30,727
: REFERENCE/DOCKET NUMBER: 19603/1521
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (716) 263-1304
: TELEFAX: (716) 263-1600
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 17 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: US-09-030-270A-7

Query Match 19.0%; Score 111; DB 2; Length 344;
Best Local Similarity 25.8%; Pred. No. 7.7e-05;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

QY 3 SIGNFNFCNIGLQTM-----GIGPOHEDSSQSPSNGSE-----COOLDQ 42
DB 58 SAGSGTMTGNTPAKDKNANAGANDPSKPSQAPSAKNTGVNDDANNDDPMQLAQ 117
QY 43 LLAMEITMKN-----LAQSSGSDANDECC--NQPPQNGQEGSLPTQMLNQLVPMQMN- 94
DB 118 LLEEDIVKLKALAHMQPGQHDKNGVNGANGAGGAGGQGLAEALQELQTLAQGGGG 177
QY 95 -----QCGAGGGGGGVNSLGN 113
DB 178 AGAGGAGGCGGGAGGAGGCGGGAGGAGGAGCGGN 212 ✓

RESULT 4
: Sequence 7, Application US/08984207
: Patent No. 6235974
: GENERAL INFORMATION:
: APPLICANT: Qiu, Dewen
: APPLICANT:

```

Mon Jul 1 09:28:18 2002

us-09-829-124-2.ra1

Page 3

```

;
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,207
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/033,230
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1201
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-984-207-7

Query Match      19.0%; Score 111; DB 4; Length 344;
Best Local Similarity 25.8%; Pred. No. 7.7e-05;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

Oy 3 SIGNNFSNIGNLQTM-----GIGPOQHESSQSPSAGSE-----QQLDQ 42
Db 58 SAGGNTGTGNAPAKGNAGANDPSKNDPSKQAPQANKTGNVDANNQDPQALQ 117
Oy 43 LLAMFTMM-----LQSQSGSDANQECG--NEQPONGQEGSLPTQMLQIMVLMQKN- 94
Db 118 LLEDLVLLKALHMQPGGNDKNGVGGANGAGAGGQGLAEALQETQILQALQGGG 177
Oy 95 -----QGGAGMGGGGSVNSLGN 113
Db 178 AGAGAGGGVGGAGAGAGGAGAGGAGAGAGAGGAGAGGAGAGGAGAGGAG 212

RESULT 5
US-09-013-587-7
; Sequence 7, Application US/09013587
; Patent No. 6277814
; GENERAL INFORMATION:
; APPLICANT: Qiu, Deven
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: ENHANCEMENT OF GROWTH IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.

```

```

;
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,587
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,048
; FILING DATE: 27-JAN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-09-013-587-7

Query Match      19.0%; Score 111; DB 4; Length 344;
Best Local Similarity 25.8%; Pred. No. 7.7e-05;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

Oy 3 SIGNNFSNIGNLQTM-----GIGPOQHESSQSPSAGSE-----QQLDQ 42
Db 58 SAGGNTGTGNAPAKGNAGANDPSKNDPSKQAPQANKTGNVDANNQDPQALQ 117
Oy 43 LLAMFTMM-----LQSQSGSDANQECG--NEQPONGQEGSLPTQMLQIMVLMQKN- 94
Db 118 LLEDLVLLKALHMQPGGNDKNGVGGANGAGAGGQGLAEALQETQILQALQGGG 177
Oy 95 -----QGGAGMGGGGSVNSLGN 113
Db 178 AGAGAGGGVGGAGAGAGGAGAGGAGAGAGGAGAGGAGAGGAGAGGAGAGG 212

RESULT 6
PCT-US96-08819-7
; Sequence 7, Application PC/TUS9608819
; GENERAL INFORMATION:
; APPLICANT: Corneil Research Foundation, Inc.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/08819
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/475,775

```

```
; FILING DATE: 07-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/10051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 7:
; LENGTH: 344 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
PCT-US96-08619-7

Query Match          19.04; Score 111; DB 5; Length 344;
Best Local Similarity 25.84; Pred. No. 7.9e-05;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

Oy 3 SIGNFNSTGNLQTM-----GIGPOHEDSSQSPSAGSE-----QQLDQ 42
Db 58 SAGTGTCTGAFKAGNAGNANDPSKNDPSKQAPQASANKTGNVDANNDPMQALMQ 117
Oy 43 LLMFIMHM-----LQSSGSDANOEG--NEQPONGQSGLSPLTOMLQIMVIMOLMKN- 94
Db 118 LLEDVYLLKALIMQPGNGKNGVGGANGAGGQGGGLAEALQETDILQALGGGG 177
Oy 95 -----QGGAGMGGGSVNSSLGN 113
Db 178 AGAGAGGCGVGGAGDGGSGGAGGANGADGDN 212

RESULT 7
US-08-891-254-3
; Sequence 3, Application US/08891254
; Patent No. 5776889
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: Hypersensitive Response
; TITLE OF INVENTION: Induced Resistance in Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon Hatgrave, Devans & Doyle
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent 1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/891,254
; FILING DATE: 10-JUL-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/475,775
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/10051
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-819-539-3

Query Match          19.04; Score 111; DB 2; Length 385;
Best Local Similarity 31.54; Pred. No. 8.9e-05;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

Oy 5 GNNFSNIGNL--OTMGI-GPOHEDSSQSPS-AGSPQOLQLIAMFIMMQLQSS--O 56
Db 118 GNNFTTSTPLQALGINTSNDSTSTGTDSSDPMOOLIKHFEIM--QSLFGD 175
Oy 57 GSDANO--ECNGNEPONGQO-----EGLSPLTOMLQIMVIMOLMKNQGGGSGVNSSL 110
Db 176 GDDGTQSSSSGKQPTGEONAYKGYDALSGLMGNGLSQLLNGNGGLAGGQGGNAGTGL 235
Oy 111 GGA 114
Db 236 DGSS 239

RESULT 8
US-08-819-539-3
; Sequence 3, Application US/08819539
; Patent No. 5859324
; GENERAL INFORMATION:
; APPLICANT: Wei, Zhong-Min
; TITLE OF INVENTION: Hypersensitive Response
; TITLE OF INVENTION: Induced Resistance in Plants
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon Hatgrave, Devans & Doyle
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/819,539
; FILING DATE: 17-MAR-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/475,775
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 14603/10050
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 385 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-819-539-3

Query Match          19.04; Score 111; DB 2; Length 385;
Best Local Similarity 31.54; Pred. No. 8.9e-05;
```


Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

OY 5 GNNFSNIGNL---QTMGI-GPOOHEDSSQSPS-AGSEQQLDQLLAFIMHMLAQSS---Q 56
DB 118 GNNFTSTNSPLDQALGINSTSDSTSDSDPMQQLKMFSEIM--QSLFGD 175
OY 57 GSDANO--ECGNEOPONGQ---BELSPLTOMLQIWMQJNONGGAGGCGGSYNSSL 110
DB 176 GQDGTQSSSGKOPTGEGNAYKGVTDALSLGMLGSLQLLGGGGLGGGAGTGL 235
OY 111 GGNA 114
DB 236 DGSS 239

RESULT 9

PCT-US93-06243-2
SEQUENCE 2, Application PC/TUS9306243
GENERAL INFORMATION:
APPLICANT: Zhong-Min Wei, David W. Bauer, Steven V.
TITLE OF INVENTION: Elicitor of the Hypersensitive Response in Plants
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Yahwak & Associates
STREET: 25 Skytop Drive
CITY: Trumbull
STATE: Connecticut
COUNTRY: U.S.A.
ZIP: 06611
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: Macintosh
OPERATING SYSTEM: MS-DOS
SOFTWARE: Microsoft Word 4.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/06243
FILING DATE: 19930630
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 907,935
FILING DATE: 01 JUL 1992
ATTORNEY/AGENT INFORMATION:
NAME: George M. Yahwak
REGISTRATION NUMBER: 26,824
REFERENCE/DOCKET NUMBER: CRF D-1172
TELEPHONE: (203)268-1951
TELEFAX: (203)268-1951
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
PCT-US93-06243-2

Query Match 19.0%; Score 111; DB 5; Length 385;
Best Local Similarity 31.5%; Pred. No. 8.9e-05;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

OY 5 GNNFSNIGNL---QTMGI-GPOOHEDSSQSPS-AGSEQQLDQLLAFIMHMLAQSS---Q 56
DB 118 GNNFTSTNSPLDQALGINSTSDSTSDSDPMQQLKMFSEIM--QSLFGD 175
OY 57 GSDANO--ECGNEOPONGQ---BELSPLTOMLQIWMQJNONGGAGGCGGSYNSSL 110
DB 176 GQDGTQSSSGKOPTGEGNAYKGVTDALSLGMLGSLQLLGGGGLGGGAGTGL 235
OY 111 GGNA 114
DB 236 DGSS 239

RESULT 11

US-08-200-724A-2
SEQUENCE 2, Application US/08200724A
PATENT NO. 5849868
GENERAL INFORMATION:
APPLICANT: Wei, Zhong-Min
APPLICANT: Bauer, David W.
APPLICANT: Beer, Steven V.
APPLICANT: Collmer, Alan

DB 236 DGSS 239

RESULT 10

PCT-US96-08819-3
SEQUENCE 3, Application PC/TUS9608819
GENERAL INFORMATION:
APPLICANT: Cornell Research Foundation, Inc.
TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
BY RESISTANCE IN PLANTS
NUMBER OF SEQUENCES: 9
CORRESPONDENCE ADDRESS:
ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
STREET: Clinton Square, P.O. Box 1051
CITY: Rochester
STATE: New York
COUNTRY: U.S.A.
ZIP: 14603
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US96/08819
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/475,775
FILING DATE: 07 JUN 1995
ATTORNEY/AGENT INFORMATION:
NAME: Goldman, Michael
REGISTRATION NUMBER: 10,727
REFERENCE/DOCKET NUMBER: 19603/10051
TELEPHONE: (716) 263-1304
TELEFAX: (716) 263-1600
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 385 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: protein
PCT-US96-08819-3

Query Match 19.0%; Score 111; DB 5; Length 385;
Best Local Similarity 31.5%; Pred. No. 8.9e-05;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

OY 5 GNNFSNIGNL---QTMGI-GPOOHEDSSQSPS-AGSEQQLDQLLAFIMHMLAQSS---Q 56
DB 118 GNNFTSTNSPLDQALGINSTSDSTSDSDPMQQLKMFSEIM--QSLFGD 175
OY 57 GSDANO--ECGNEOPONGQ---BELSPLTOMLQIWMQJNONGGAGGCGGSYNSSL 110
DB 176 GQDGTQSSSGKOPTGEGNAYKGVTDALSLGMLGSLQLLGGGGLGGGAGTGL 235
OY 111 GGNA 114
DB 236 DGSS 239

APPLICANT: He, Sheng-Yang
 APPLICANT: Laby, Ron J.
 TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE
 IN PLANTS
 NUMBER OF SEQUENCES: 5
 CORRESPONDENCE ADDRESS:
 ADDRESS: Milton, Margrave, Devans & Doyle
 STREET: Clinton Square
 CITY: Rochester
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 14603

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:
 APPLICATION NUMBER: 09/200,724A
 FILING DATE: 23-FEB-1994

CLASSIFICATION: 510

ATTORNEY/AGENT INFORMATION:
 NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/10030

TELECOMMUNICATION INFORMATION:
 TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:
 LENGTH: 403 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-200-724A-2

Query Match 19.0%; Score 111; DB 2; Length 403;
 Best Local Similarity 31.5%; Pred. No. 9.4e-05;
 Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

Oy 5 GNFSNIGNL---OTMGI-GPOQHESSQSPS-AGSEQQLQQLIAMFIMHMQQS---Q 56

Db 118 GNVITSTNSPLDQALGINSTSDSTSGTSDSDSDPMQOLLKMFSEIM--QSLFPGD 175

Oy 57 GSDANO--ECGNBPQNGCOQ---EGLSPLTQMQLQVIMQMGAGMGCGGSVNSL 110

Db 176 GQDTGCGSSGGKOPTGEONAYKKGVTDLGSLGMLGSLGGLGGLGGLGGLGGLGGL 235

Oy 111 GGN A 114

Db 236 DGSS 239

RESULT 12

US-09-030-270A-3

Sequence 3, Application US/09030270A

Patent No. 5977060

GENERAL INFORMATION:

APPLICANT: Zitter, Thomas A.

APPLICANT: Wei, Zhong-Min

TITLE OF INVENTION: INSECT CONTROL WITH A

HYPERSENSITIVE RESPONSE ELICITOR

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESS: Milton, Margrave, Devans & Doyle LLP

STREET: P.O. Box 1051, Clinton Square

CITY: Rochester

STATE: New York

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: 09/200,724A

FILING DATE: 23-FEB-1994

CLASSIFICATION: 510

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/10030

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 403 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-08-200-724A-2

MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/030,270A
 FILING DATE:

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/039,226

FILING DATE: 28-FEB-1997

ATTORNEY/AGENT INFORMATION:

NAME: Goldman, Michael L.

REGISTRATION NUMBER: 30,727

REFERENCE/DOCKET NUMBER: 19603/1521

TELECOMMUNICATION INFORMATION:

TELEPHONE: (716) 263-1304

TELEFAX: (716) 263-1600

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 403 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: Protein

US-09-030-270A-3

Query Match 19.0%; Score 111; DB 2; Length 403;
 Best Local Similarity 31.5%; Pred. No. 9.4e-05;
 Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

Oy 5 GNFSNIGNL---OTMGI-GPOQHESSQSPS-AGSEQQLQQLIAMFIMHMQQS---Q 56

Db 118 GNVITSTNSPLDQALGINSTSDSTSGTSDSDSDPMQOLLKMFSEIM--QSLFPGD 175

Oy 57 GSDANO--ECGNBPQNGCOQ---EGLSPLTQMQLQVIMQMGAGMGCGGSVNSL 110

Db 176 GQDTGCGSSGGKOPTGEONAYKKGVTDLGSLGMLGSLGGLGGLGGLGGLGGLGGL 235

Oy 111 GGN A 114

Db 236 DGSS 239

RESULT 13

US-08-851-376A-2

Sequence 2, Application US/08851376A

Patent No. 6187172

GENERAL INFORMATION:

APPLICANT: Beer, Steven V.

APPLICANT: Wei, Zhong-Min

APPLICANT: Bauer, David W.

APPLICANT: Collmer, Alan

APPLICANT: He, Sheng-Yang

APPLICANT: Laby, Ron

TITLE OF INVENTION: ELICITOR OF THE HYPERSENSITIVE RESPONSE

IN PLANTS

NUMBER OF SEQUENCES: 5

CORRESPONDENCE ADDRESS:

ADDRESS: Milton, Margrave, Devans & Doyle LLP

STREET: Clinton Square, P.O. Box 1051

CITY: Rochester

STATE: NY

COUNTRY: U.S.A.

ZIP: 14603

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/851,376A

```

; FILING DATE: 05-MAY-1997
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/700,724
; FILING DATE: 23-FEB-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/10035
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-851-376A-2

Query Match 19.0%; Score 111; DB 4; Length 403;
Best Local Similarity 31.5%; Pred. No. 9.4e-05;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

Oy 5 GNFSNIGML--QTMGI-GPOHEDSSQSPS-AGSPQQLDOLLAMFIMMLQQS---Q 56
Db 118 GNTTSTNSPLDQALGINSTSGTSTGSDSDPQOOLLKAFSEIN--QLSRGD 175

Oy 57 GSDANQ--ECGNEQPONGQO---EGLSPLTQMLQIYVLMQNGGAGGCGGSYNSSL 110
Db 176 GQDCTGCGSSGGKQPTGEQNAKKGVTALSGLNGLSQLLNGGLGGGCGGNAGTGL 235

Oy 111 GGNA 114
Db 236 DGSS 239

RESULT 14
US-08-984-207-3
; Sequence 3, Application US/08984207
; Patent No. 6235974
; GENERAL INFORMATION:
; APPLICANT: Qiu, Dewen
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: HYPERSENSITIVE RESPONSE INDUCED
; TITLE OF INVENTION: RESISTANCE IN PLANTS BY SEED TREATMENT
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: P.O. Box 1051, Clinton Square
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/984,207
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE: 05-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1201

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-984-207-3

Query Match 19.0%; Score 111; DB 4; Length 403;
Best Local Similarity 31.5%; Pred. No. 9.4e-05;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;

Oy 5 GNFSNIGML--QTMGI-GPOHEDSSQSPS-AGSPQQLDOLLAMFIMMLQQS---Q 56
Db 118 GNTTSTNSPLDQALGINSTSGTSTGSDSDPQOOLLKAFSEIN--QLSRGD 175

Oy 57 GSDANQ--ECGNEQPONGQO---EGLSPLTQMLQIYVLMQNGGAGGCGGSYNSSL 110
Db 176 GQDCTGCGSSGGKQPTGEQNAKKGVTALSGLNGLSQLLNGGLGGGCGGNAGTGL 235

Oy 111 GGNA 114
Db 236 DGSS 239

RESULT 15
US-09-013-587-3
; Sequence 3, Application US/09013587
; Patent No. 6277814
; GENERAL INFORMATION:
; APPLICANT: Qiu, Dewen
; APPLICANT: Wei, Zhong-Min
; APPLICANT: Beer, Steven V.
; TITLE OF INVENTION: ENHANCEMENT OF GROWTH IN PLANTS
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon, Hargrave, Devans & Doyle LLP
; STREET: Clinton Square, P.O. Box 1051
; CITY: Rochester
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 14603
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/013,587
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/036,048
; FILING DATE: 27-MAR-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Goldman, Michael L.
; REGISTRATION NUMBER: 30,727
; REFERENCE/DOCKET NUMBER: 19603/1501
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (716) 263-1304
; TELEFAX: (716) 263-1600
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 403 amino acids
; TYPE: amino acid
; STRANDEDNESS: linear
; TOPOLOGY: linear
; MOLECULE TYPE: protein

```

US-09-013-587-3

Query Match 19.0%; Score 111; DB 4; Length 403;
Best Local Similarity 31.5%; Pred. No. 9.4e-05;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;
OY 5 GNVFSNIGL---QTMGI-GPOHEDSSQSPS-AGSEODLDOLLAMFIMMIDQS---Q 56
DB 118 GNVFTSTNSPLDQALGINSTSONDSTISGTSTSDSDPQQLAMFSEIM--QSLFGD 175
OY 57 GSDANO--ECNEOPONGQ----EGLSPLTOMLMOIVHOLMONGGAGMCGGGSVNSL 110
DB 176 GQDGTGSSGCGKQPTGEBONAYKGVTDALSGLMNGLSQLLGNGLGCGGCGNAGTGL 235
OY 111 GNA 114
DB 236 DGS 239

Search completed: June 28, 2002, 10:21:52
Job time: 319 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2002, 10:16:58 ; Search time 14.36 Seconds
(without alignments) 762.826 Million cell updates/sec

Title: US-09-829-124-2

Perfect score: 584

perfect score: 384
Sequence: 1 MDSIGNFNSIGNLQTMGIG.....QQGAGMGGGGSVNSSLGNA 114

Scoring table:

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

```
Maximum DB seq length: 0
Maximum DB seq length: 2000000000
```

Post-processing: Minimum Match 0%

Cost Processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

```

PRN/I:
1: p1r1:

```

```

1: p1r1: *
2: p1r2: *

```

3: p1r3: *

4: p1r4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	111	19.0	344	2	A51707	1	PopAI protein - Pa
2	111	19.0	403	2	T08471	2	harpin - Erwinia a
3	96.3	16.5	272	2	G75548	3	hypothetical protein
4	96.3	16.5	196	2	A31068	4	neurogenic locus m
5	95.4	15.4	334	2	G86395	5	hypothetical protein
6	90.5	15.9	348	2	T22330	6	hypothetical protein
7	90.5	15.9	782	2	T63248	7	hypothetical protein
8	90.5	15.5	748	2	T25238	8	hypothetical protein
9	90	15.4	707	2	T13638	9	hypothetical protein
10	90	15.4	3190	2	G73749	10	CRMS-binding protein
11	89.5	15.3	597	1	S37849	11	cDNA intrastand cr
12	88	15.1	2414	2	A52127	12	transcription adap
13	87	14.9	663	2	G21912	13	hypothetical protein
14	87	14.9	728	2	T21913	14	brocote-WPI-O
15	86	14.7	794	2	T10519	15	pre-pro-leugmin
16	85.5	14.6	195	2	G83170	16	hypothetical protein
17	85.5	14.6	521	2	A40232	17	elav protein - fru
18	85.5	14.6	641	2	PC7036	18	heat shock protein
19	84.5	14.5	915	2	G24577	19	ovarian protein -
20	84	14.4	447	2	T18447	20	HrpW protein - Erw
21	84	14.4	353	1	A35615	21	transcription init
22	83	14.2	426	2	T51861	22	hypothetical protein
23	83	14.2	1656	2	G53316	23	probable protein k
24	83	14.2	1552	2	KC3938	24	gene mastermind pr
25	82.5	14.1	791	2	T19235	25	probable Athila re
26	82.5	14.1	1145	2	T19235	26	transcription acti
27	82.5	14.1	1480	2	S48440	27	poly(A)-specific r
28	82	14.0	696	2	S71899	28	immune factor 1. t
29	81	13.9	660	2	A34266	29	glutelin high mole

ALIGNMENTS

```

RESULT 1
S41707
Popal protein - Pseudomonas solanacearum
C:Species: Pseudomonas solanacearum
C:Date: 13-Jan-1999 #sequence_revision 13-Jan-1995 #text_change 08-Oct-1999
C:Accession: S41707; 598621
E:Arlat, M.; van Gijsegem, F.; Huet, J.C.; Pernollet, J.C.; Boucher, C.A.
EMBO J. 13, 543-553, 1994
A:Title: Popal, a protein which induces a hypersensitivity-like response on specific
A:Reference number: S41707; MUID:94148001
A:Accession: S41707
A:Status: preliminary
A:Accession type: DNA
A:Residue type: DNA
A:Cross-references: EMBL:227092; NID:9414909; PIDN:CAAB1615.1; PID:9414910

Query Match 19.0%; Score 111; DB 2; Length 344;
Best Local Similarity 25.8%; Pred. No. 0.0016;
Matches 40; Conservative 17; Mismatches 54; Indels 44; Gaps 5;

QY 3 SIGNNFSNIGNTQM-----GIQPOHDSQSQSFASGE-----QQIDQ 42
58 SAGNTGTGTPAKPKAGNAGNAGNPDSKNDPSQAPSKNTGTVDDANNDDPMQALMQ 117
QY 43 LLANTIMW-----LAQSGSDANDEG--NEQPOQOQEGELPTQMLHQTIVQLMNH- 94
DB 118 LLEDVLKLAALHMQPGNDKNGVGGNGAGNGAGGQGGQAEALAEIQLAQGGGG 177
QY 95 -----QQGAGGGGGGSSVNSLGN 113
DB 178 AGAGGAGGGVGGAGGADGGSGAGGAGGAGGADGGN 212

RESULT 2
T08471
harpin - Erwinia amylovora
C:Species: Erwinia amylovora
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 11-Jun-1999
C:Accession: T08471;
E:Laby, R.J.; Kim, J.F.; Beer, S.V.
Submitted to the EMBL data Library, March 1999
A:Reference number: 216433
A:Accession: T08471
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-403 <IAB>
A:Cross-references: EMBL:M92994; NID:94309677; PID:94309678
A:Genetic source: strain 321
A:Gene: hrpN

```

Query Match 19.0%; Score 111; DB 2; Length 403;
Best Local Similarity 31.5%; Pred. No. 0.0019;
Matches 39; Conservative 21; Mismatches 48; Indels 16; Gaps 7;
QY 5 GNFFSNGNL--QTGNI-GPQHHSSQSPS-AGSEQLDQLLAFIMMLQOQS---Q 56
DB 118 GNTTSTNSPLDQALGINSTSONDSTGTSTSDSPNQQLKMFSEIM--OSLPD 175
OY 57 GSDAQ---ECGNEQPNQOQ---EGLSPITQMLQIVYMLQNOGAGGSGGSVNSL 110
DB 176 QDGTGSSGCKQPTGEGNATKGVTDLSGLNGLSQLLNGSLGGGGAGTGL 235
OY 111 GGVA 114
DB 236 DGSS 239
RESULT 3
C75348
Hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
R:Accession: C75348 J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;
M.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Oterbach, T.; Zalewski, C.; Ma
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
S:Title: Genome sequence of the radioreistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: C75348
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-272 <WIL>
A:Cross-references: GB:AE001882; GB:AE000513; NID:96457865; PIDN:AAE09792.1; PID:9645787
A:Experimental source: strain R1
A:Genetic: DBR201
A:Map position: 1
Query Match 16.5%; Score 96.5; DB 2; Length 272;
Best Local Similarity 25.5%; Pred. No. 0.033;
Matches 39; Conservative 15; Mismatches 36; Indels 63; Gaps 7;
OY 1 MDSIGNFNIGNQTMGIPQGHSSQSPSAGSEQLDQLLAFIMMLQOQSQSDA 60
DB 41 LGAMTRATPGGLDALASAVGDDSA-----MLGFGQGTGDM 80
QY 61 NQ-----ECGNEQPNQOQ---EGLSPITQMLQIVY-----Q 90
DB 81 NEDQKILGHVFGGQAAANNIGDAGIDP--QNMGLMAYAPILSYLSNRSRAQGG 138
QY 91 LMQNGG-----AGMG--GGSVNSLGG 112
DB 139 MCGNNGGSGGNSLPGNGGGLGGLGSLGG 171
RESULT 4
A33106
neurogenic locus mmm protein - fruit fly (Drosophila melanogaster)
A:Alternative names: mastermind protein
A:Title: neurogenic locus mmm protein
C:Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 21-Jul-2000
R:Accession: A36391; A33106; S135
R:Smoller, D.; Friedel, C.; Schmid, A.; Bettler, D.; Lam, L.; Yedvobnick, B.
Genes Dev. 4, 1688-1700, 1990
A:Title: The Drosophila neurogenic locus mastermind encodes a nuclear protein unusually
A:Reference number: A36391; MUID:91065516
A:Accession: A36391
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1596 <SHO>

A:Cross-references: GB:X54251; NID:98203; PIDN:CAA38152.1; PID:98204
A:Note: strain Canton S
C:Genetic: DBR201
A:Genes: FlyBase:mam
A:Cross-references: flyBase:FBgn0002643
Query Match 16.5%; Score 96.5; DB 2; Length 1596;
Best Local Similarity 28.8%; Pred. No. 0.22;
Matches 32; Conservative 7; Mismatches 39; Indels 33; Gaps 4;
OY 22 QHREDSQSPSAGSEQLDQLLAFIMMLQOQSQSDANQCGNEQPNQOQEGLSPLT 81
DB 1428 QQQQQQQHMGFGAANNMOQLLQO-----QSSGGG-----GNMAMOMOMTSMH-MT 1474
OY 82 QMLQIVYMLQIVY-----QGGAGGSGGSVNSLGG 112
DB 1475 QTQQITMQQQQPVQSTTTTTHQOQOQWQMGPGGGGGGPGCSANNNGG 1525
RESULT 5
G86385
Hypothetical protein F237.4 [Imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Dec-2001
R:Accession: G86385
R:Heideberg, J.A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alu
A:Authors: Salzman, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tall
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719
A:Accession: G86385
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-538 <STO>
A:Cross-references: GB:AE005172; NID:910092326; PIDN:AGI2737.1; GSPDB:GN00141
C:Genetic: DBR201
A:Map position: 1
Query Match 16.4%; Score 96; DB 2; Length 534;
Best Local Similarity 34.1%; Pred. No. 0.077;
Matches 31; Conservative 4; Mismatches 44; Indels 12; Gaps 2;
OY 22 QHREDSQSPSAGSEQLDQLLAFIMMLQOQSQSDANQCGNEQPNQOQEGLSPLT 81
DB 436 QHQQQQQQQQQQQQQQHQLTQHHH-----QOQQQNSPLAQHQOOTSPLNOMQOOTSPLN 491
OY 82 QMLQIVYMLQIVY-----QGGAGGSGGG 104
DB 492 QNQQQQQPOQWVNGGQAPGAPGSGGQGGG 522
RESULT 6
T22330
Hypothetical protein F47A4.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 18-Feb-2000
R:Accession: T22330
R:Hortmore, B.
Submitted to the EMBL Data Library, June 1995
A:Accession: T22330
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-3498 <WIL>
A:Cross-references: EMBL:Z49888; PIDN:CAA90064.1; GSPDB:GN00028; CESP:F47A4.2

C:Genetics:
A:Cross-references: FlyBase:FBgn0015624
A:Map position: X
C:Superfamily: bromodomain homology
F:1723-1780/Domain: bromodomain homology <BRO>

Query Match 15 48; Score 90; DB 2; Length 3190;
Pair Local Similarity 30.1%; Fried No 2;
Matches 28; Conservative 35; Indels 24; Gaps 3;
QY 15 QTGIGGIP--QHEDSSQGSAGSGSQQLDLAMFIMHMLQSGSDANQCGNEDPQNG 72
DB /2875 QGANGSPGNPQGGGQGGGQGGGQGGGQGGGQGGGQGGGQGGGQGGGQGGG 2926
OY 73 QGSELSPLTQMLQIYVQLMQLNGGAGNGGGGS 105
DB 2927 QG-----VNLQMQGGGQGGGPGGS 2945

RESULT 11
S37849
DNA intrastrand crosslink recognition protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein YKL032c; protein YKL245; transcription factor ORD1
C:Species: Saccharomyces cerevisiae
C:Date: 10-Sep-1999 sequence revision
C:Accession: S37849; S37853; S41668; S39002; A57739
C:Submitted to: the Protein Sequence Database, March 1994
A:Reference number: S37832
A:Accession: S37849
A:Molecule type: DNA
A:Residues: 1-597 <R1>
A:Cross-references: EMBL:X71632; NID:G486033; PIDN:CAA81867.1; PID:G486034; GSPDB:GN0001
A:Experimental source: strain S288C
A:Experimental source: B.; Skala, J.; van Dyck, L.; Tettelin, H.; Goffeau, A.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S37851
A:Accession: S37851
A:Molecule type: DNA
A:Residues: 1-597 <R2>
A:Cross-references: EMBL:X71632; NID:G486033; PIDN:CAA81867.1; PID:G486034; GSPDB:GN0001
A:Experimental source: strain S288C
A:Experimental source: B.; Skala, J.; van Dyck, L.; Tettelin, H.; Goffeau, A.
submitted to the Protein Sequence Database, March 1994
A:Reference number: S41667
A:Accession: S41667
A:Molecule type: DNA
A:Residues: 1-597 <R3>
A:Cross-references: EMBL:X71632; NID:G505182; PIDN:CAB37853.1; PID:G4467991
A:Experimental source: strain S288C
A:Experimental source: J.; Kellert, P.J.; Lippard, S.J.
submitted to the EMBL Data Library, July 1993
A:Description: IKR1, a yeast protein that binds to platinated DNA and confers sensitivity to cisplatin
A:Reference number: S39002
A:Accession: S39002
A:Molecule type: DNA
A:Residues: 1-72-92, 'F', 94-107, 113-206, 'T', 208-219, 'TT', 222-597 <BRO>
A:Cross-references: EMBL:U16590; NID:g311108; PIDN:AAA02859.1; PID:g311109
A:Experimental source: J.; Bilanchone, V.M.; Cumsky, M.G.
Proc. Natl. Acad. Sci. U.S.A. 91, 7345-7349, 1994
A:Title: The ORD1 gene encodes a transcription factor involved in oxygen regulation and
A:Reference number: A57739; MUID:94316692
A:Accession: A57739
A:Experimental source: strain S288C
A:Experimental source: B.; Skala, J.; van Dyck, L.; Tettelin, H.; Goffeau, A.
submitted to the Protein Sequence Database, March 1994
A:Reference number: GB:L16900
A:Accession: GB:L16900
A:Molecule type: DNA
A:Residues: 1-107, 113-597 <LAW>
A:Cross-references: GB:L16900
C:Genetics:
A:Gene: SGD:IXR1; ORD1; MIPS:YKL032C

C:Genetics:
A:Gene: FlyBase:br
A:Cross-references: FlyBase:FBgn0000210
C:Superfamily: POZ domain homology
C:Keywords: zinc finger
F:18-119/domain: POZ domain homology <POZ>

C;Cross references: Flybase:FBgn00000210
C;Superfamily: POZ domain homology
C;Keywords: zinc finger
E;F:18-119/Domain: POZ domain homology <POZ>

```

Query Match          14.9%  Score 87:  DB 2:  Length 728;
Best Local Similarity 28.3%  Pred. No. 0.83;
Matches 32; Conservative 10; Mismatches 51; Indels 20; Gaps 4;

Dy      5  GNNFSIGNLQTMIGIPQOQHEDSSQSPGASQEQDQLLAMPINMLQSQSGSDANQRC 64
Db      511  GSN-----GGLSNLTALLPQQQQQQQQQQQQQQQQQQQ-----QQQQQQ 564

Dy      65  GREQPPQNGQQEGLSPLE-----TQMLAQTMVQLMQLQGAGMGGGGVSSVSLGNN 113
Db      565  GQQQQPPNGIACSPFKANTPTTTTQQQQQAAYVMAAASASASTSGSANSLSNNS 617

```

Experimental source: tissue_type seed
Cross-references: EMBL:Y09117; NID:e1008798; PID:e284333
Keywords: seed; storage protein

Query Match 14.7%; Score 86; DB 2; Length 794;
Best Local Similarity 29.5%; Pred. NO. 1.1;
Matches 36; Conservative 15; Mismatches 53; Indels 18; Gaps 5;

a

7 NFSNIGNLQTMGI-----GPOHEDSSQQSPSA--GSEQQLDQLLAMFIMMLQQSQGS 58
 :|:| :|:| :|:| :|:| :|:| :|:| :|:|
b 194 DFANAQNQLDRNVRRFFLAGNOOHVAQLGGGSSMKGESAGCNVNLAFGLDLILQSGLI 253

[illegible]

111 66'112

[illegible]

b 312 GG 313

Search completed: June 20, 2002, 10:22:21

Job time: 323 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: June 28, 2002, 10:20:53 ; Search time 22.09 Seconds
(without alignments)
892.776 Million cell updates/sec

Title: US-09-829-124-2
Perfect score: 584
Sequence: 1 MOSIGNNFSNIGLQTMGIG.....OGAGNGGGGVSNSLAGNA 114

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL19.*
1: sp.archaea.*
2: sp.bacteria.*
3: sp.fungi.*
4: sp.human.*
5: sp.invertebrate.*
6: sp.mammal.*
7: sp.mhc.*
8: sp.organelle.*
9: sp.phage.*
10: sp.plant.*
11: sp.protist.*
12: sp.virus.*
13: sp.vertebrate.*
14: sp.unclassified.*
15: sp.virus.*
16: sp.bacteriap.*
17: sp.archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	189.5	32.4	139	2 O9KW32	O9KW32 xanthomonas
2	189	32.4	143	2 O9Z1J5	O9Z1J5 xanthomonas
3	114	19.5	330	2 O9WKK0	O9WKK0 burkholderi
4	110.5	18.9	838	5 O9YOC9	O9YOC9 dictyostell
5	110	18.8	344	2 O9RAS0	O9RAS0 burkholderi
6	104.5	19.2	318	5 O9WOP3	O9WOP3 drosophila
7	104.5	17.2	318	2 O9WOP3	O9WOP3 burkholderi
8	100.5	17.2	1173	2 O9RAT0	O9RAT0 burkholderi
9	96.5	16.5	272	16 O9RAT5	O9RAT5 deinothecus
10	96.5	16.5	1366	5 O9RAT5	O9RAT5 drosophila
11	96.5	16.5	1594	5 O9V6W7	O9V6W7 drosophila
12	96	16.4	534	10 O9C5M3	O9C5M3 xanthomonas
13	92.5	15.8	3498	5 Q20497	Q20497 caenorhabdi
14	91.5	15.7	424	2 O87264	O87264 pseudomonas
15	91	15.6	624	5 O94486	O94486 dictyostell
16	90.5	15.5	748	10 O65609	O65609 arabidopsis

17	90.5	15.5	748	10 O9MOK8	O9MOK8 arabidopsis
18	90	15.4	707	5 Q23191	Q23191 caenorhabdi
19	90	15.4	748	5 O95TP4	O95TP4 drosophila
20	90	15.4	2328	5 O9VM47	O9VM47 drosophila
21	90	15.4	2531	5 O9GPH4	O9GPH4 drosophila
22	90	15.4	3190	5 O01368	O01368 drosophila
23	90	15.4	3275	5 O9M321	O9M321 drosophila
24	88	15.2	1483	3 P79025	P79025 emericeila
25	88	15.2	1483	3 P79025	P79025 emericeila
26	87	14.9	1778	5 O9V6W7	O9V6W7 drosophila
27	87	14.9	2035	5 O9G519	O9G519 drosophila
28	86	14.7	794	10 P93560	P93560 arabidopsis
29	85.5	14.6	195	16 O9HXK1	O9HXK1 pseudomonas
30	85.5	14.6	642	3 O9UVW0	O9UVW0 rhizopus st
31	85.5	14.6	2994	5 O95ZG5	O95ZG5 dictyostell
32	85	14.6	147	5 O9V543	O9V543 drosophila
33	85	14.6	365	2 O9KH45	O9KH45 erwinia her
34	85	14.6	713	10 O9SDM2	O9SDM2 triticum ae
35	85	14.6	713	10 O94IK8	O94IK8 secale cere
36	85	14.6	713	10 O94IK8	O94IK8 secale cere
37	85	14.6	713	10 O94IK8	O94IK8 secale cere
38	84.5	14.5	707	5 O9V6W7	O9V6W7 drosophila
39	84.5	14.5	915	5 O24433	O24433 drosophila
40	84.5	14.5	915	5 O24433	O24433 drosophila
41	84.5	14.5	915	5 O9VCN7	O9VCN7 drosophila
42	84.5	14.5	2653	5 O25253	O25253 lucilia cup
43	84	14.4	447	2 O54508	O54508 erwinia amy
44	84	14.4	447	2 O9LAW2	O9LAW2 erwinia amy
45	84	14.4	926	5 O9M3G1	O9M3G1 drosophila

ALIGNMENTS

RESULT 1
ID O9KW32 PRELIMINARY: PRT: 139 AA.
AC O9KW32
DT 01-OCT-2000 (TREMREL. 15, Created)
DT 01-OCT-2000 (TREMREL. 15, Last sequence update)
DT 01-DEC-2001 (TREMREL. 19, Last annotation update)
DE HPAL HOMOLOG.
OS Xanthomonas oryzae pv. oryzae.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX Xanthomonas
RN Xanthomonas
RX MEDLINE-2130248; PubMed-11410350;
RP SEQUENCE FROM N.A.
RC STRAIN-MAFF 311018;
RX MEDLINE-2130248; PubMed-11410350;
RA Ochiai H., Inoue Y., Hasebe A., Kaku H.;
RT "Construction and characterization of a Xanthomonas oryzae pv. oryzae
RT Bacterial artificial chromosome library."
RL FEMS Microbiol. Lett. 200:59-65(2001).
DR EMBL: AB045311; BAB07848.1; --
SQ SEQUENCE 139 AA: 13727 MW: D05857FD8E403898 CRC64;

Query Match 32.4%; Score 189.5; DB 2; Length 139;
Best Local Similarity 40.7%; Pred. No. 3.5e-13;
Matches 50; Conservative 12; Mismatches 48; Indels 13; Gaps 5;
QY 1 MOSIGNNFSNIGLQTMGIGPOQRED--SSQGSFSAQ--SSQQLDQLLAFMMMLQOS---55
DB 1 MNSLNTQFG--GSTSNLQVGFSDPTFGSNGGNGISSEKQLDQLLQSSKNA 58
QY 56 -----OGSDANGECNEDPQNGQDGLSPTQNLMTQVLMQNGGAGGGGVSNSL 110
DB 59 BEKGQGGGNGGCGGNSQQAGQONGGSPFTQNLMLHIVGISTQANGGAGGGG-FOGG 117
QY 111 GGN 113
DB 118 GGD 120

```

RESULT 2
Q921J5 PRELIMINARY: PRT: 143 AA.
AC Q921J5
DT 01-MAY-1999 (TRENBLrel. 10, Created)
DT 01-MAY-1999 (TRENBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TRENBLrel. 19, Last annotation update)
GN HPA1
OS Xanthomonas oryzae pv. oryzae.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OX NCBI_TaxID=64187;
RN [1]
RP STRAIN=PX086;
RX MEDLINE=20179737; PubMed=10714988;
RA Zhu W., MacBanua M.M., White F.F.;
RT Identification of two novel hrp-associated genes in the hrp gene
RT cluster of Xanthomonas oryzae pv. oryzae.*;
DB EMBL:AB026107; AAC59211;
SQ SEQUENCE 143 AA; 31987 MW; 053BC594CA241E5C CRC64;

Query Match 32.4%; Score 189; DB 2; Length 143;
Best Local Similarity 41.0%; Pred. No. 4.1e-13;
Matches 50; Conservative 11; Mismatches 49; Indels 12; Gaps 4;

QY 1 NDSIGNFSGNIGLQTMGTPQGHED-SSQQSPFAG-SEQQLDQLLAFIMMLQAS--- 55
Db 1 NWSLNTQFG-GSTNLQVSPQDTTFGSGQNGQISEKQLDQLLQALQSRA 58
QY 56 -----QGSNDANFQNDPQNGQSGLSPLTQMLQIVMLQNGGAGGGSNSL 110
Db 59 ECKGKGDDGGGGGNSQAGQNGSPFTUHLHIVEILQNGGGAGGGGFGGFG 118
QY 111 GG 112
Db 119 GG 120

RESULT 3
Q9WKKO PRELIMINARY: PRT: 330 AA.
AC Q9WKKO
DT 01-NOV-1999 (TRENBLrel. 12, Created)
DT 01-NOV-1999 (TRENBLrel. 12, Last sequence update)
DT 01-NOV-1999 (TRENBLrel. 12, Last annotation update)
DE POPAI PROTEIN.
GN POPAI.
OS Burkholderia solanacearum (Pseudomonas solanacearum).
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OX NCBI_TaxID=305;
RN [1]
RP "Popai ORF of Ralstonia solanacearum".
RT Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL:AB026239; BAA7270.1;
SQ SEQUENCE 330 AA; 31997 MW; 8EC4C2D9E319DA5 CRC64;

Query Match 19.5%; Score 114; DB 2; Length 330;
Best Local Similarity 28.4%; Pred. No. 0.00018;
Matches 40; Conservative 15; Mismatches 56; Indels 30; Gaps 5;

QY 3 SIGNFSGNGL-----QTMGICPQHR-----RDSQSQSPFAGSEQQLDQ 42
Db 111

```

RL	1	Mol. Microbiol.	36:261-277(2000).
RR	2	SEQUENCE FROM N.A.	
RR	3	STRAIN-GM1000;	
RR	4	MEDLINE-95349395;	
RR	5	Van Gijsegem F., Gough C.L., Zischek C., Genlin S., Niqueux B.,	
RR	6	Arlat M., Barberis P., German S., Estérel J., Boucher C.A.;	
RR	7	"The hrp gene locus of <i>Pseudomonas solanaceae</i> encodes a positive regulator of	
RR	8	III secretion system, encodes eight proteins related to components of	
RR	9	the flagellar biogenesis complex.";	
RR	10	Mol. Microbiol. 15:1095-1114(1995).	
RR	11	[3]	
RR	12	SEQUENCE FROM N.A.	
RR	13	STRAIN-GM1000;	
RR	14	MEDLINE-93125128; PubMed-1479894;	
RR	15	Genlin S., Gough C.L., Zischek C., Boucher C.A.;	
RR	16	"Evidence that the hrpB gene encodes a positive regulator of	
RR	17	pathogenicity genes from <i>Pseudomonas solanaceae</i> .";	
RR	18	Mol. Microbiol. 6:3065-3076(1992).	
RR	19	[4]	
RR	20	SEQUENCE FROM N.A.	
RR	21	STRAIN-GM1000;	
RR	22	MEDLINE-93113006; PubMed-1472716;	
RR	23	Gough C.L., Genlin S., Zischek C., Boucher C.A.;	
RR	24	"hrp genes of <i>Pseudomonas solanaceae</i> are homologous to pathogenicity	
RR	25	determinants of animal pathogenic bacteria and are conserved among	
RR	26	plant pathogenic bacteria.";	
RR	27	Mol. Plant Microbe Interact. 5:384-389(1992).	
RR	28	[5]	
RR	29	SEQUENCE FROM N.A.	
RR	30	STRAIN-GM1000;	
RR	31	MEDLINE-93302711; PubMed-8316211;	
RR	32	Gough C.L., Genlin S., Lopes V., Boucher C.A.;	
RR	33	"Homology between the hrpO protein of <i>Pseudomonas solanaceae</i> and	
RR	34	bacterial proteins implicated in a signal peptide-independent secretion	
RR	35	mechanism.";	
RR	36	Mol. Gen. Genet. 239:378-392(1993).	
RR	37	[6]	
RR	38	SEQUENCE FROM N.A.	
RR	39	STRAIN-GM1000;	
RR	40	MEDLINE-96347139; PubMed-8736546;	
RR	41	Bogdanov A.J., Beer S.V., Bonas U., Boucher C.A., Collmer A.,	
RR	42	Caplin D.L., Cornelis G.R., Huang H.C., Hutcheson S.W.,	
RR	43	Panopoulos N.J., Van Gijsegem F.;	
RR	44	"Unified nomenclature for broadly conserved hrp genes of	
RR	45	phytopathogenic bacteria.";	
RR	46	Mol. Microbiol. 20:681-683(1996).	
RR	47	[7]	
RR	48	SEQUENCE FROM N.A.	
RR	49	STRAIN-GM1000;	
RR	50	MEDLINE-94148001; PubMed-8313899;	
RR	51	Arlat M., Van Gijsegem F., Huet J., Pernollet J., Boucher C.A.;	
RR	52	"Specific protein which induces a hypersensitivity-like response on	
RR	53	specific <i>Pseudomonas</i> genotypes, is secreted via the hrp pathway of	
RR	54	<i>Pseudomonas solanaceae</i> .";	
RR	55	EMBO J. 13:543-553(1994).	
RR	56	[8]	
RR	57	SEQUENCE FROM N.A.	
RR	58	STRAIN-GM1000;	
RR	59	MEDLINE-94247349; PubMed-8190064;	
RR	60	Genlin S., Boucher C.A.;	
RR	61	"A superfamily of proteins involved in different secretion pathways in	
RR	62	gram-negative bacteria: modular structure and specificity of N-terminal	
RR	63	domain.";	
RR	64	Mol. Gen. Genet. 243:112-118(1994).	
RR	65	EMBL: AJ245811; CAB58262.1.	
RR	66	SEQUENCE 344 AA; 33152 MW; 4247C5A0C3280907 CRC64;	

Query Match 18.8%; Score 110; DB 2; Length 344;
Best Local Similarity 25.8%; Pred. No. 0.0005;
Matches 40; Conservative 16; Mismatches 55; Indels 44; Gaps 5;

QY	3	STGNFSGNIGLQTN	-----GIGPOHDSQSPAGSG-----	QOLDQ 42
Db	58	SAGNTGNTGNPAKD	GNAGANDYR ⁵ SKID ⁵ SKGSG ⁵ POSANTGYVDNANDNDPQALMQ	117
QY	43	LLAMFIMMM	-----LQSSGSDANEGC-----NEOPNSQSGSLSPMLQIMQIVHMLGN-94	
Db	118	LLLEVLKLAALH	QPGNDKNGVGGANGAGGSGGLAERAEQTEIILAJGGGG 177	
QY	95	-----	-----OGAGNGGSGSVNSIGN 113	
Db	178	AGAGGAGGGVGGAGG	AGDGGSGAGGAGGAGDGN 212	
RESULT	6			
ID	QWOP3	PRELIMINARY;	PRT; 318 AA.	
AC	QWOP3			
DT	01-MAY-2000 (TMBLrel. 13; Created)			
DT	01-MAY-2000 (TMBLrel. 13; Last sequence update)			
DT	01-MAY-2000 (TMBLrel. 13; Last annotation update)			
DE	CG17183 PROTEIN (SDS560P).			
GN	CG17183			
OS	Drosophila melanogaster (Fruit fly).			
OC	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
OC	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;			
OC	Ephydroidea; Drosophilidae; Drosophila.			
OX	NCBI_Taxid=7227;			
RP	[1]			
RP	SEQUENCE FROM N.A.			
RC	SPRAIN-BERKELEY;			
RC	MEDLINE=20196006; PubMed=10731132;			
RA	Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,			
RA	Georgieva B., Gish W.B., Hamilton W., Hoskins R.A., Galle R.F.,			
RA	Sutton R.C., Wortman J.R., Yeung M.D., Holt R.A., Chen X.,			
RA	Brandon R.C., Rogers Y.-H.C., Blazer V.G., Chang M., Pfeiffer B.D.,			
RA	Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,			
RA	Abri J.F., Agbayani A., An H.-J., Andrews-FannKoch C., Baldwin D.,			
RA	Ballew R.M., Basu P.A., Baxendale J., Bayraktaroglu L., Beasley E.M.,			
RA	Beeson K.Y., Benson P.V., Berman B.P., Bhargava D., Bolshakov S.,			
RA	Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,			
RA	Kurtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,			
RA	Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,			
RA	deCapelis B., Decker A., Deng Z., Mays A.D., Dew I., Dietz S.M.,			
RA	DiGiuseppe J., Dugan-Rocha S., Dunks B.C., Dunn P.,			
RA	Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasner K.,			
RA	Glodzik A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M., Houck J.,			
RA	Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,			
RA	Hoslin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,			
RA	Jatelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,			
RA	Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,			
RA	Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,			
RA	Liu X., Mattet B., McIntosh T.C., McLeod M.P., McPherson D.,			
RA	Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,			
RA	Mout R.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,			
RA	Palazzo M., Pittman S., Proctor S., Raskin D., Riedel J.M.,			
RA	Palazzo M., Pittman S., Proctor S., Raskin D., Riedel J.M.,			
RA	Reinert K., Remington K., Saunders R.C., Scheeler P., Sheng H.,			
RA	Shue B.C., Siden-Klamis I., Simpson M., Skupski M.P., Smith T.,			
RA	Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,			
RA	Swirski R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,			
RA	Wang Z.-Y., Wassman D.A., Weinstein G.C., Weissenbach J.,			
RA	Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,			
RA	Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,			
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,			
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;			
RT	The genome sequence of <i>Drosophila melanogaster</i> ;			
RT	Science 287:2185-2195;(2000).			
RL	SEQUENCE FROM N.A.			
RA	Stapleton M., Brockstein P., Hong L., Agbayani A., Carlson J.,			
RA	Chamce W., Chavez D., Dorsett Y., Farnham D., Friesen J.,			

```

00 1 MDSNGNNGNMGTCGQPOHEDSQSPAGSQQLQLAMFTMMHLQSQGSDA
DB   : : : : : : : : : : : : : : : : : : : : : : : : : : : :
41 LGAATNATGATGGDLASAGHNDQSA-----MIGGCGGCTGCTDWM
      Matches 39; Conservative 15; Mismatches 36; Indels 63; Gaps 0
Best Match 16.5%; Score 96.5; DB 16; Length 272;
Query Local Similarity 25.5%; Pred. No. 0.012;

```

OY 61 NO-----ECNEQFON-----GQCEGLSPLTOMLQIVW-----Q 90
DB 81 NEGQKILGHVFGQCOAAANAIGRAGIDP--QMAQILMAVPLILSYLSNRSBAGGQ 138
OY 91 LMONOGG-----AGMGG--GGSVNSLGG 112
DB 139 MGNMGQSGGMSLPCGMGGGLGGILGG 171

RESULT 10
OYV698 PRELIMINARY; PRT: 1366 AA.

AC QV698
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DE MAN PROTEIN.
GN MAN OR CG8118.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RX NCBI_TaxID=7277;
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthites P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Abril J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Beasley E.M.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis J.M., Cawley S., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Fostler C., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,
RA Hostin D., Houston K.A., Howland T.J., Karpis S., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA McQuinn G., Milshina N.V., Mobarry C., McLeod M.P., McPherson D.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
RA Shue B.C., Sidel-Kimos I., Stapleton M., Strong R., Sun E.,
RA Spier E., Spradling A.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT Science 287:2185-2195(2000).
RT ENBL: AE001816; AAF58300.1;
DR FlyBase: FBgn002643; man.
SQ SEQUENCE 1366 AA; 141456 MW; CC10F8D5E23F3631 CRC64;

Query Match 16.5%; Score 96.5; DB 5; Length 1366;
Best Local Similarity 28.8%; Pred. No. 0.073;

Matches 32; Conservative 7; Mismatches 39; Indels 33; Gaps 4;
OY 22 QHEDSSQSFSGSQQDLOOLAMFPMHLOQSGSDANOCNRPQNGQSLT 81
DB 1198 QQQQQQHHGFGAANMNOOLLOO-----QSGGG-----GNASQMDTSHH MT 1244
OY 82 QMLQIVMQLNQ-----QSGAGMGCGGSVNSLGG 112
DB 1245 QTCOOITMOOQGVOSTTTTTHOOQMMQMGCGGCGGSGSANNNGG 1295

RESULT 11
OYV697 PRELIMINARY; PRT: 1594 AA.

AC QV697
DT 01-MAY-2000 (TREMREL. 13, Created)
DT 01-MAY-2000 (TREMREL. 13, Last sequence update)
DE MAN PROTEIN.
GN MAN OR CG8118.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
RX NCBI_TaxID=7277;
RP SEQUENCE FROM N.A.
RC STRAIN=BERKELEY.
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananthites P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champagne M., Pfeiffer B.D.,
RA Abril J.F., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Balow R.M., Basu A., Baxendale J., Bayraktaroglu L., Baldwin D.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Beasley E.M.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,
RA Burtis J.M., Cawley S., Cadieu E., Center A., Chandra I.,
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Fostler C., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houch J.,
RA Hostin D., Houston K.A., Howland T.J., Karpis S., Kravitz S., Kulp D., Lai Z.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA McQuinn G., Milshina N.V., Mobarry C., McLeod M.P., McPherson D.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Smith T.,
RA Shue B.C., Sidel-Kimos I., Stapleton M., Strong R., Sun E.,
RA Spier E., Spradling A.C., Turner R., Venter E., Wang A.H., Wang X.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Zheng X.H., Zhong F.N., Zhou W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Zeng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RT Science 287:2185-2195(2000).
RT ENBL: AE001816; AAF58300.1;
DR FlyBase: FBgn002643; man.
SQ SEQUENCE 1594 AA; 167263 MW; D4236DAN26P70D092 CRC64;

Query Match 16.5%; Score 96.5; DB 5; Length 1594;


```

DR EMBL: AF005321; AAC62526.1; -.
DR EMBL: AF23004; AAF71503.1; -.
SQ SEQUENCE 424 AA; 42910 MW; 7AF4ED059BE79D2E CRC64;

Query Watch 15.78; Score 91.5; DB 2; Length 424;
Best Local Similarity 27.88; Pred. No. 0.06; 27; Indels 27; Gaps
Matches 27; Conservative 16; Mismatches 16;

QY 26 DSSQDQPSFAG-----SEQDLQLLAFMTMMHQLQSGSDNAGCGNEQPGNGQBSGLSPL 80
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
61 DSTVONPODASKPNDOSNIKALISALIMSLQMLTNSKNKQDNGEQP-----DSQAP- 114
QY 81 TQMLMIVMQLMONGGAG-----MGGGGVSWSSLGG 112
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
115 -----FQNNGGIATPSADSGGGGTDPATGGG 140

RESULT 15
794486 PRELIMINARY; PRT: 624 AA.
AC 034486
DC 034486
AT 01-FEB-1997 (T-EMBLrel. 02, Created)
DT 01-FEB-1997 (T-EMBLrel. 02, Last sequence update)
DT 01-FEB-1997 (T-EMBLrel. 02, Last annotation update)
DE ORF DG1040 (FRAGMENT)
DE Dictyostelium discoideum (slime mold)
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OC NCBI_TaxID:44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN:AX4;
RA Loomis W.F.;
RL Submitted (AUG-1996) to the EMBL/GenBank/DBJ databases.
RL EMBL: U66913; AAB07544.1; -.
FT NON-TER 624 624
SQ SEQUENCE 624 AA; 70445 MW; 50343A91ECB0CB37 CRC64;

Query Watch 15.58; Score 91; DB 5; Length 624;
Best Local Similarity 28.18; Pred. No. 0.12;
Matches 27; Conservative 10; Mismatches 37; Indels 22; Gaps
Matches 74

QY 22 QHEDSSQSPSAGSEQLDQLLAFMTMMHQLQSGSDNAGCGNEQPGNGQO----- 74
DB ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
298 QQQNNQQQQNNQQQQQQQQQQQ-----QQQQRQQQQQQQQNQQQQQQNQHYSEYIK 347
QY 75 -----BELSLPITQMLQIVMQLMONGGAGMGGGGSVNS 108
DB 348 YLQGISPNSLALS-----PLVNDNQSSSGSGGGGINS 380.

```

Search completed: June 28, 2002, 10:23:23
Job time: 150 sec

Result No.	Query No.	Score	%			ID	Description
			Match	Length	DB		
1	1	111	19.0	403	1	HRPN_ERWAM	Q01099 erwinia amy
2	2	110	18.8	344	1	POPA_RALSO	Q9rbso ralstonia s
3	3	96.5	16.5	1596	1	MAW_DROME	P21519 drosophila s
4	4	89.5	15.3	597	1	IXRL_YEAST	P33417 saccharomyc
5	5	88	15.1	2414	1	P300_HUMAN	Q09472 homo sapien
6	6	87	14.9	727	1	BRCL_HROME	Q01295 drosophila s
7	7	86.5	14.8	519	1	ELAV_DROVI	P23241 drosophila s
8	8	85	14.6	675	1	EGR_AEDAE	P49880 aedes aegypt
9	9	83	14.2	353	1	TF2D_DROME	P20227 drosophila s
10	10	82.5	14.1	1480	1	PANL_YEAST	P32521 saccharomyc
11	11	81	13.9	660	1	GLT3_WHEAT	P08488 triticum ae
12	12	79	13.5	416	1	R23B_MOUSE	P54728 mus musculu
13	13	79	13.5	648	1	GLT0_WHEAT	P10387 triticum ae
14	14	78.5	13.4	512	1	FUS_BOVIN	Q28009 bos taurus
15	15	78	13.4	255	1	LP6L_EIEMTE	P15714 elmeria ten
16	16	78	13.4	1018	1	YK26_CAEEL	P34333 caenorhabdi
17	17	77	13.2	685	1	YYCA_BACSU	P37483 bacillus su
18	18	76.5	13.1	405	1	YCC8_YEAST	P25367 saccharomyc
19	19	76.5	13.1	829	1	E74A_DROME	P20105 drosophila s
20	20	76.5	13.1	883	1	E74B_DROME	P11536 drosophila s
21	21	76	13.0	834	1	5HTA_DROME	P28285 drosophila s
22	22	75	12.8	363	1	TOB1_MOUSE	Q61471 mus musculu
23	23	75	12.8	526	1	FUS_HUMAN	P35637 homo sapien
24	24	74.5	12.8	463	1	K10_DROME	P33468 drosophila s
25	25	74.5	12.8	648	1	KAPC_DICDI	P34099 dictyostell
26	26	74.5	12.8	1516	1	NCO2_XENLA	Q9W705 xenopus lae
27	27	74.5	12.8	1845	1	R23B_HUMAN	Q9u136 homo sapien
28	28	74	12.7	409	1	R24B_HUMAN	P54727 homo sapien
29	29	73.5	12.6	738	1	YKF4_YEAST	P35732 saccharomyc
30	30	73.5	12.6	1043	1	FTFL_DROME	P33244 drosophila s
31	31	73.5	12.6	2038	1	FSH_DROME	P13709 drosophila s
32	32	73	12.5	269	1	BDNF_XIPMA	Q02193 xiphophorus
33	33	73	12.5	593	1	K1CJ_HUMAN	P13645 homo sapien

```

176 GDDTQCGSSGGKOPTGEQNAAYKGVTDALSGMLGNGLSQLCLNGLGSGGGCGGNAGTGL 233
QY 111 GGNA 114
DB 236 DGSS 239

RESULT 2
POPA_RALSO
ID POPA_RALSO STANDARD; PRT; 344 AA.
AC Q9RBS0: Q52544; Q9WXX0: Q9REY0;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE POPa1 protein [Contains: POPa2 protein; POPa3 protein].
GN POPA OR RSP0877 OR RS01648.
OS Ralstonia solanacearum (Pseudomonas solanacearum).
OC Plasmid megaplasmid.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
OC Ralstonia.
OX NCBI_TaxID=305;
OX [1]
RN RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RX MEDLINE=94148001; PubMed=8313899;
RA Ariat M., Van Gijsegem F., Huet J.-C., Pernollet J.-C., Boucher C.A.;
RT "Popa1, a protein which induces a hypersensitivity-like response on
RT specific Petunia genotypes, is secreted via the Hrp pathway of
RT Pseudomonas solanacearum.";
RL EMBO J. 13:543-553(1994).
RN [2]
RN RP SEQUENCE FROM N.A.
RC STRAIN=GM11000;
RA Salanoubat M., Genin S., Artiguenave F., Gouzy J., Mangenot S.,
RA Arlat M., Billault A., Bottier P., Camus J.-C., Cattolico L.,
RA Chandler M., Choine N., Claudel-Renard C., Cunnac S., Demange N.,
RA Gaspin C., Lavie M., Moisan A., Robert C., Saurin W., Schliex T.,
RA Sigur P., Thebaud P., Whalen M., Wincker P., Levy M.,
RA Weissenbach J., Boucher C.A.;
RT "Genome sequence of the plant pathogen Ralstonia solanacearum.";
RN Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN RP SEQUENCE FROM N.A.
RC STRAIN=8107;
RA Hikichi Y., Ikegami M., Okuno T.;
RN Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
RN [4]
RN RP SEQUENCE FROM N.A.
RC STRAIN=OEI-1;
RA Hikich Y., Kanda A., Hasegawa H., Okuno T.;
RT "Popa1 in Ralstonia solanacearum OEI-1 pathogenic to tobacco.";
RN Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
-1- FUNCTION: Act as a specific hypersensitive response (HR) elicitor
CC Has activity on tobacco (non-host plant) and petunia but is
CC without activity on tomato (host plant).
CC -1- SUBCELLULAR LOCATION: Secreted.
CC -1- PTM: POPa2 and POPa3 are produced from POPa1 by stepwise removal
CC of N-terminal amino acids
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabora-
CC between the Swiss Institute of Bioinformatics and the EMBL outstatin
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in
CC modified and this statement is not removed. Usage by and for commu-
CC entities requires a license agreement (See http://www.isb-sib.ch/annou-
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AJ245811; CAB59262.1; -
CC DR EMBL; AL646081; CAD18028.1; -
CC DR EMBL; AB026629; BAA77270.1; -
CC DR EMBL; AB032747; BAA84679.1; -
CC Hypersensitive response; Plasmid; Complete proteome.

```


[illegible]

ID	ECR_AEDAE	STANDARD;	PRT;	675 AA.
----	-----------	-----------	------	---------

Query Match	14.6%	Score 85;	DB 1;	Length 675;
Best Local Similarity	29.6%	Pred. No. 0.47;		
Matches 29: Conservative	10;	Mismatches 39;	Indels 20;	Gaps 4;

QY 22 QHEDSSQSPPSAGSEQQLDQLLAWFIMMLQQSQG-SDANQECGNPQNQGQ-----E 71
Db 580 QMHSHTGSSTSSSSSSSS-----SSNGSGNSGSSNSSNQHPHPHQ 626

QY 76 GLSPLTQMLOIWMQIMO-NOGGAGMGGGGVNSSLG 112
Db 627 QLTPNQOQHOOHOLQOVHANGSGSGGNNSSSG 664

RESULT 9

ID TF2D_DROME STANDARD; PRT; 353 AA.

AC P02027; O3WZD8;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-FEB-1991 (Rel. 17, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Transcription Initiation factor TFIID (TATA-box factor) (TATA
sequence-binding protein) (TBP) (TATA-box binding protein).

GN TBP OR TFIID OR BTF1 OR CG9874.
OS Drosophila melanogaster (fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OX Ephydroidea; Drosophilidae; Drosophila.
NCBI_TaxId=7227;

[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=90304877; PubMed=2194666;
EX Hoey T., Dynlacht B.D., Peterson M.G., Pugh B.F., Tjian R.;
RA "Isolation and characterization of the Drosophila gene encoding the
RT TATA box binding protein, TFIID.";
RL Cell 61:1179-1186(1990).
RN [2]
RN SEQUENCE FROM N.A.
RP MEDLINE=91067664; PubMed=2123550;
RX Munich M., Iida C.T., Horikoshi M., Roeder R.G., Parker C.S.;
RA *cDNA clone encoding Drosophila transcription factor TFIID.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:9148-9152(1990).
RN [3]
RN SEQUENCE FROM N.A.
RP Lira-Devito L.M., Burke T.W., Kadonaga J.T.;
RL Submitted (JUN-1994) to the EMBL/GenBank/DBJ databases.
RN [4]
RN SEQUENCE FROM N.A.
RP STRAIN=OREGON-R;
RC Lee K., Oh Y., Yoon J., Cho N., Baek K.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [5]
RN SEQUENCE FROM N.A.
RP STRAIN=BKKELEY;
RC MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Vandeil M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballwey R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotter P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablo S., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Foster C., Gabrielson A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ikegawa C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

DT 01-FEB-1995 (Rel. 31, Last sequence update)
 DE 01-NOV-1997 (Rel. 35, Last annotation update)
 GN PAN1 protein.
 GN PAN1 OR MIP3 OR MIP3 OR YIR006C OR YIB6C.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
 NCBI_TaxID=4932;
 [1]
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 320-344; 352-375 AND 899-906.
 RX MEDLINE=92405166; PubMed=1339314;
 RA Sachs A.B., Deardorff J.A.;
 RT "Translation initiation requires the PAB-dependent poly(A)
 RL ribonuclease in yeast.";
 RL Cell 70:961-973(1992).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=5288C / AB972;
 RA Barrell B.G., Badcock K., Bankier A.T., Bowden S., Brown D.,
 RA Churcher C.M., Connor R., Copsey T., Dear S., Devlin K., Fraser A.,
 RA Gentles S., Hamlyn N., Horsnell T.S., Hunt S., Jagels K., Jones M.,
 RA Louis E., Lye G., Moule S., Moule T., Odell C., Pearson D.,
 RA Rajandream M.A., Riles L., Rowley N., Skelton J., Smith V.,
 RA Walsh S.V., Whitehead S.;
 RA Submitted (DEC-1994) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN SEQUENCE FROM N.A.
 RC STRAIN=5288C;
 RX MEDLINE=95282515; PubMed=7762303;
 RA Voss H., Tamames J., Teodoro C., Valencia A., Senses C., Wiemann S.,
 RA Schwager C., Zimmermann J., Sander C., Ansorge W.;
 RA "Nucleotide sequence and analysis of the centromeric region of yeast
 RT chromosome IX";
 RL Yeast 11:61-78(1995).
 CC -|- FUNCTION: NOT KNOWN.
 CC -|- PTM: THE N-TERMINUS IS BLOCKED.
 CC -|- SIMILARITY: SOME TO MAMMALIAN EPS15.
 CC -|- CAUTION: WAS ORIGINALLY THOUGHT TO BE A SUBUNIT OF PAB-DEPENDENT
 CC POLY(A)-SPECIFIC RIBONUCLEASE.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z38062; CAA86208.1;
 DR EMBL; M90688; AAA34841.1;
 DR EMBL; X79743; CAB38097.1;
 DR PIR; S30889; S30889.
 DR PIR; S48440; S48440.
 DR SGD; S0001445; PAN1.
 DR InterPro: IPR000261; EPS15_repeat.
 DR Pfam: PF00036; efband; 3.
 DR SMART; SM00054; Efb; 2.
 DR SMART; SM00027; Efb; 2.
 KW Repeat.
 FT DOMAIN 142 575 15 X 12 AA APPROXIMATE REPEATS.
 FT REPEAT 142 153 1-1.
 FT REPEAT 164 175 1-2.
 FT REPEAT 188 199 1-3.
 FT REPEAT 215 226 1-4.
 FT REPEAT 235 246 1-5.
 FT REPEAT 392 403 1-6.
 FT REPEAT 409 420 1-7.
 FT REPEAT 422 433 1-8.
 FT REPEAT 446 457 1-9.
 FT REPEAT 467 478 1-10.
 FT REPEAT 498 509 1-11.
 FT REPEAT 507 518 1-12.
 FT REPEAT 538 549 1-13.

FT REPEAT 545 556 1-14.
 FT REPEAT 564 575 1-15.
 FT DOMAIN 328 680 2 X 23 AA APPROXIMATE REPEATS.
 FT REPEAT 328 350 2-1.
 FT REPEAT 658 680 2-2.
 FT DOMAIN 1084 1125 7 X 6 AA TANDEM REPEATS.
 FT REPEAT 1084 1089 3-1.
 FT REPEAT 1090 1095 3-2.
 FT REPEAT 1096 1101 3-3.
 FT REPEAT 1102 1107 3-4.
 FT REPEAT 1108 1113 3-5.
 FT REPEAT 1114 1119 3-6.
 FT REPEAT 1120 1125 3-7.
 FT DOMAIN 1315 1377 8 X 6 AA APPROXIMATE REPEATS.
 FT REPEAT 1315 1320 4-1.
 FT REPEAT 1321 1326 4-2.
 FT REPEAT 1327 1332 4-3.
 FT REPEAT 1340 1345 4-4.
 FT REPEAT 1345 1350 4-5.
 FT REPEAT 1355 1360 4-6.
 FT REPEAT 1361 1366 4-7.
 FT REPEAT 1372 1377 4-8.
 FT DOMAIN 13 22 POLY-GLN.
 FT DOMAIN 29 34 POLY-GLN.
 FT DOMAIN 98 106 POLY-GLN.
 FT DOMAIN 1400 1406 POLY-PRO.
 FT DOMAIN 1452 1455 POLY-GLU.
 FT DOMAIN 1474 1480 POLY-PRO.
 FT CONFLICT 235 235 P -> T (IN REF. 1).
 FT CONFLICT 266 273 ITAQDAK -> YICPSGKN (IN REF. 1).
 FT CONFLICT 474 487 MISSING (IN REF. 1).
 FT CONFLICT 653 657 MISSING (IN REF. 1).
 FT CONFLICT 1291 1291 A -> R (IN REF. 1).
 FT CONFLICT 1396 1480
 SQ SEQUENCE 1480 AA; 160267 MW; F3518495FF759553 CRC64;
 Query Match 14.1%; Score 82.5; DB 1; Length 1480;
 Best Local Similarity 27.6%; Pred No. 2;
 Matches 29; Conservative 9; Mismatches 44; Indels 23; Gaps 4;
 QY 15 QTMGIGQPHEDSSQQPSAGSEQQLDQLAMFMMLQQSQGSQDANQECNEQFPQ---- 70
 Db 6 QQQGGMGYQQQQQQQQQQPQNG-----FYPQQQGGSSNQPGQPQQQMA 50
 QY 71 --NCQQBGLSPLTOMLMQIVMQLMNOGGAGMGCGGGSVNSLGN 113
 Db 51 FNOFQATGIGGMPQSGFSGFSMPQ-QPQTGYNNNGN-NGSVYGN 93
 RESULT 11
 GLT3_WHEAT
 ID GLT3_WHEAT STANDARD; PRT; 660 AA.
 AC P08488;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUL-1999 (Rel. 38, Last annotation update)
 DE Glutenin, high molecular weight subunit 12 precursor.
 OS Triticum aestivum (Wheat).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Triticeae; Triticum.
 NCBI_TaxID=4565;
 [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=CV. CHINESE SPRING;
 RX MEDLINE=86041882; PubMed=3840588;
 RA Thompson R.D., Bartels D., Harberd N.P.;
 RT "Nucleotide sequence of a gene from chromosome 1D of wheat encoding a

RT HMW-glutenin subunit.
CC Nucleic Acids Res. 13:6833-6846(1985).
CC -!- FUNCTION: GLUTENINS ARE THE HIGH MOLECULAR WEIGHT SEED STORAGE
CC PROTEIN OF WHEAT ENDOSPERM. THOUGHT TO BE RESPONSIBLE FOR THE
CC VISCO-ELASTIC PROPERTY OF WHEAT DOUGH.
CC -!- SUBUNIT: DISULFIDE-BRIDGE LINKED AGGREGATES.
CC -!- MISCELLANEOUS: GLUTENINS ARE CODED BY SEVERAL GENES ON EACH OF THE
CC GROUP 1 CHROMOSOMES OF WHEAT.
CC -!- MISCELLANEOUS: THE MATURE PROTEIN IS CHARACTERIZED BY A LARGE
CC NUMBER OF WELL PRESERVED REPEATS OF THE TWO MOTIFS: GQDPGQ AND
CC GQDPGQGGQGYPTS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X03041; CAA26847.1; -
DR PIR; A24266; A24266.
DR HSSP; P01088; 1BFA.
DR InterPro; IPR001768; Cereal tryp_amy1_inh.
DR InterPro; IPR001419; Glutenin.
DR Pfam; PF00234; tryp-alpha_amy1; 1.
DR PRINTS; PR00210; GLUTENIN.
KW Seed storage protein; Repeat; Multigene family; Signal.
FT SIGNAL 1 21
FT CHAIN 22 660
FT DOMAIN 125 615
FT REPEATS.
SQ SEQUENCE 660 AA; 70868 MW; 2BFD09D8C8FC0CFF CRC64;

Query Match 13.9%; Score 81; DB 1; Length 660;
Best Local Similarity 27.4%; Pred. No. 1.1;
Matches 29; Conservative 10; Mismatches 45; Indels 22; Gaps 4;

QY 12 GNLQTMIGPQHQHEDSSQSPSAGSEQLDQLLAMFTMMMLQSQGSDANQECNEQPN 71
Db 264 GQWQSGGQGGHYPTSLQPGGQ-----GHYLSAQQPAQGGQGHYPASQQQPGQ 316

QY 72 GQDEGLSPLTQMLQIVMLQMQGG-----AGMGGGGSVNS 109
Db 317 GQQ-GHVPASQ-----QDPGCGQGHYPASQGPQGGQGGQIPAS 355

RESULT 12
R23B_MOUSE
ID R23B_MOUSE STANDARD; PRT; 416 AA.
AC P54728;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE UV excision repair protein RAD23 homolog B (MHR23B) (XP-C repair
DE complementing complex 58 kDa protein) (P58).
GN RAD23B OR MHR23B.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=BAUB/C; TISSUE=Testis;
RX MEDLINE=96403997; PubMed=8808275;
RA van der Spek P.J., Visser C.E., Hanaoka F., Smit B.,
RA Hagemeijer A., Bootsma D., Hoeijmakers J.H.J.;
RT "Cloning, comparative mapping, and RNA expression of the mouse
RT homologues of the Saccharomyces cerevisiae nucleotide excision repair
RT gene RAD23.";
RL Genomics 31:20-27(1996).
CC -!- FUNCTION: INVOLVED IN DNA EXCISION REPAIR. MAY PLAY A PART IN DNA

DAMAGE RECOGNITION AND/OR IN ALTERING CHROMATIN STRUCTURE TO
ALLOW ACCESS BY DAMAGE-PROCESSING ENZYMES.
CC -!- SUBUNIT: HETERODIMER OF A 125 kDa SUBUNIT (P125) AND OF A
CC 58 kDa SUBUNIT (P58).
CC -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC -!- SIMILARITY: CONTAINS 1 UBIQUITIN-LIKE DOMAIN.
CC -!- SIMILARITY: CONTAINS 2 UBA DOMAINS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; X92411; CAA63146.1; -
DR HSSP; P54725; 1DV0.
DR MGI; MGI:105128; Rad23b.
DR InterPro; IPR000449; UBA.
DR InterPro; IPR000626; Ubiquitin.
DR Pfam; PF00627; UBA; 2.
DR Pfam; PF00240; ubiquitin; 1.
DR SMART; SM00165; UBA; 2.
DR SMART; SM00213; UBO; 1.
DR PROSITE; PS00553; UBIQUITIN_2; 1.
KW DNA damage; DNA repair; Nuclear protein.
FT DOMAIN 1 79 UBIQUITIN-LIKE.
FT DOMAIN 255 261 POLY-ALA.
FT DOMAIN 262 270 POLY-THR.
FT DOMAIN 336 355 POLY-GLY.
SQ SEQUENCE 416 AA; 43516 MW; 13E0245A6D892205 CRC64;

Query Match 13.5%; Score 79; DB 1; Length 416;
Best Local Similarity 30.8%; Pred. No. 1.1;
Matches 28; Conservative 9; Mismatches 30; Indels 24; Gaps 3;

QY 22 QQHEDSSQSPSAGSEQLDQLLAMFTMMMLQSQGSDANQECNEQPNQGGGLSPLT 81
Db 286 QQMROITTOQNPSTL-----LPALL-----QQIGRENPLQLQ--ISQHQ 321

QY 82 QMLMOIVMQLMONQAGMGGSVNSSLGG 112
Db 322 EHFQMLNEPQVEAGGGGGGGGGGGGG 352

RESULT 13
GLTO_WHEAT
ID GLTO_WHEAT STANDARD; PRT; 648 AA.
AC P10387;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Glutenin, high molecular weight subunit D110 precursor.
GN GLU-D1-2B.
OS Triticum aestivum (Wheat).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4565;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. CHEYENNE;
RX MEDLINE=89098419; PubMed=2563152;
RA Anderson O.D., Greene F.C., Yip R.E., Halford N.G., Shewry P.R.,
RA Malpica-Romero J.M.;
RT "Nucleotide sequences of the two high-molecular-weight glutenin genes
RT from the D-genome of a hexaploid bread wheat, Triticum aestivum L. cv
RT Cheyenne.";
RL Nucleic Acids Res. 17:461-462(1989).
CC -!- FUNCTION: GLUTENINS ARE THE HIGH MOLECULAR WEIGHT SEED STORAGE
CC PROTEIN OF WHEAT ENDOSPERM. THOUGHT TO BE RESPONSIBLE FOR THE

RC TISSUE-Sporozoite;
RX MEDLINE-30348718; PubMed-2200963;
RA KO C., Smith C.K. II, McDonnell M.;
RT "Identification and characterization of a target antigen of a
monoclonal antibody directed against *Eimeria tenella* merozoites.";
RL Mol. Biochem. Parasitol. 41:53-64(1990).
CC -!- FUNCTION: UNKNOWN. THE GLN-RICH TANDEM REPEATS MAY BE IMPORTANT
FOR AN UNKNOWN ASPECT OF THE PARASITIC LIFE CYCLE. MAY BE AN
CC IMPORTANT IMMUNOGEN.
CC -!- SUBUNIT: MAY BE COVALENTLY LINKED BY DISULFIDE BONDS TO OTHER
POLYPEPTIDES TO FORM THE 80 KDa ANTIGEN.
CC -!- DEVELOPMENTAL STAGE: PRESENT IN ALL STAGES THROUGHOUT THE
SPORULATION OF THE OOCYSTS AND IN THE SPOOROZOITES FOLLOWING
CC EXCYSTATION.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M30933; AAA29079.1;
DR PIR; A60637; A60637.
KW Antigen; Sporozoite; Repeat; Sporulation.
FT NON_TER 1
FT DOMAIN 18 210 12 X APPROXIMATE TANDEM REPEATS, GLN-
RICH.
FT REPEAT 18 48 1.
FT REPEAT 49 57 2.
FT REPEAT 58 65 3.
FT REPEAT 66 78 4.
FT REPEAT 79 90 5.
FT REPEAT 91 103 6.
FT REPEAT 104 140 7.
FT REPEAT 141 152 8.
FT REPEAT 153 164 9.
FT REPEAT 165 172 10.
FT REPEAT 173 192 11.
FT REPEAT 193 210 12.
FT NON_TER 255 255
SQ SEQUENCE 255 AA; 31267 MW; 8C5E6005FFFC2DB3 CRC64;

Query Match 13.4%; Score 78; DB 1; Length 255;
Best Local Similarity 32.6%; Pred. No. 0.79;
Matches 30; Conservative 6; Mismatches 48; Indels 8; Gaps 2;
QY 22 QHEDSSQSPSAGSEQQDQLLAFIMMLOQSGSDANQECGNQPNQGGQGLSPLT 81
DB 144 000000000WPEQPEQQ0000WPEQ-----QQQQWSDNQQQQAGQWQAQQQQQWPPQQ 197
QY 82 QMLQIIVMLQNOGAGMGCGGGSVNSLGGN 113
DB 198 QQPQQ--QQQQQQQDGLGPDGIVVPLGSS 227

Search completed: June 28, 2002, 10:22:46
Job time: 168 sec